Neural mechanisms underlying voluntary action control in the human brain

Nima Khalighinejad

Institute of Cognitive Neuroscience

University College London

PhD Thesis

2017

Declaration

I, Nima Khalighinejad, confirm that the work presented in this thesis is my own. Any information derived from other sources is fully cited and referenced in the thesis.

Nima Khalighinejad

Abstract

The capacity for voluntary action lies at the heart of human nature. Healthy adults typically feel that they can exercise voluntary control over their actions, and thus change their physical and social environment. This feeling of autonomous control over one's own action is a foundational concept for most human societies. However, neuromechanistic models of this ubiquitous experience remain unclear. This thesis explores the neural mechanisms underlying voluntary action control in the human brain and the experiences associated with it such as the experience of agency. After operationalizing voluntary action in a novel behavioural paradigm, we show that self-initiated actions are preceded by a specific preparatory process in the brain. Later experiments suggest that the experience of agency might be a metacognitive readout of this precursor process: a study of a patient with anarchic hand syndrome shows that precursor processes for voluntary action contribute to the sense of agency over outcomes of action. We then provide new causal evidence that the experiences of voluntary action could be influenced by modulating specific parts of the brain that may host these precursor processes such as parietal and frontal areas. Finally, we show that by pairing voluntary actions of one hand with involuntary movements of the other hand, key aspects of agency experience can transfer from voluntary to involuntary movements after repeated association. This later finding suggests that the experiences of voluntary action are not hardwired, but could be acquired through associative learning. This thesis concludes by proposing that a dedicated system in the human brain contributes to the preparation and execution of self-initiated voluntary actions and the characteristic subjective experiences associated with it.

Acknowledgements

I am immensely grateful to my supervisor, Prof Patrick Haggard, for all his support during my PhD. In the last three years, not only did he teach me to design and conduct scientific research but how to think critically, to make effective arguments, to disseminate my research and to establish collaborations and work as part of a large research team, as well as independently. I enormously enjoyed being part of the Action and Body group.

All of the works within this thesis are the result of valuable scientific input from my supervisor, numerous collaborators within and outside UCL, and financial support from the ERC advanced grant HUMVOL.

For chapter 2, I wish to thank Dr Aaron Schurger for helping me with computational modelling, interpretation of the data, and for allowing me to visit his lab. I wish to thank Dr Andrea Desantis for his contribution in designing the experimental paradigm, and Leor Zmigrod for assisting in collecting the data.

For chapter 3, I wish to thank Prof Elisabetta Ladavas and Dr Caterina Bertini for allowing me to work in their lab and their effort in recruiting the anarchic hand patient for the study. Special thanks go to patient TP for her time and contribution. I would also like to thank Anamika Kunnumpurath who helped with data collection from healthy participants.

For chapter 4, I wish to thank Jack Booth and Ksenia Vinogradova for their help in collecting the data. The material in this chapter has been partly published as “Khalighinejad, N., & Haggard, P. (2015). Modulating human sense of agency with non-invasive brain stimulation. *Cortex*. 69, 93–103”.

For chapter 5, I wish to thank Steven Di Costa for his contribution in designing and conducting some of the experiments and Julian Mutz who helped with data collection. The material in this chapter has been partly published as “Khalighinejad, N., Di Costa, S., &

Haggard, P. (2016). Action selection processes in dorsolateral prefrontal cortex contribute to sense of agency: a meta-analysis of tDCS studies of ‘intentional binding’. Brain Stimulation. 9(3), 372-379”.

For chapter 6, I wish to thank Dr Rui Costa and Dr Rick Adams for their helpful advice and discussion. I would like to give special thanks to Dr Rui Costa for allowing me to work in his lab. It was an unforgettable experience. The material in this chapter has been partly published as “Khalighinejad, N., & Haggard, P. (2016). Extending experiences of voluntary action by association. Proceedings of the National Academy of Sciences. 201521223”.

Finally, I wish to thank my wife, Dr Talitha de Vries, for her endless love and support.

Contents

List of figures

List of tables

Chapter 1

Introduction and literature review

1.1. Introduction

Actions are the essential content of our daily life. Without actions we could not obtain food, we could not run away from dangers, and we could not reproduce or obtain any other goal in life. They are so essential to our existence that some argue our brains have evolved to serve the demands of action rather than cognition (Llinas, 2002). Traditionally, human actions have been categorised into reactions: actions that happen in response to some kind of external stimuli, sometimes automatically, and voluntary actions that are generated endogenously at a time of our own free choice, are free from external constraints and are associated with specific subjective experiences ('it just feels like that it was my decision'). Most of our daily actions fall somewhere on the spectrum between fully reactive to fully volitional actions. For example pressing the brake pedal in response to a red light when we are driving a car has elements of both types of actions: it is voluntary because the action belongs to us and we may consciously decide to slow down in good time. At the same time it is a reaction, because we initiate the action in response to an external signal.

The capacity for voluntary action is a critical component of what defines us as humans. We typically feel fully in control of our actions. We are able to make changes in our environment. We hold people responsible for their actions. We punish them for making the wrong action. Interestingly, even though we experience these feelings in our everyday life and have a strong sense of freedom, the capacity in freely choosing between different courses of action has been challenged since earliest times. Epicurus is known to be the first thinker to recognise the 'free will problem' when he said: "Necessity destroys responsibility and chance is inconstant; whereas our own actions are autonomous, and it is to them that praise and blame naturally attach" (Reale, 1985). Even though the 'hard problems' of free will have concerned philosophers for the last 25 centuries, the recent advancements in neurosciences have opened new opportunities to illuminate empirical aspects of this problem. Importantly, scientific investigations on the mechanisms underlying voluntary action control in the human

brain not only reshape current understanding of how the human brain works, but could also have major implications for scientific understanding of society generally.

1.2. What is volition?

1.2.1. Action as the final output

As Sherrington pointed out in one of his lectures “to move things is all that mankind can do, for such the sole executant is muscle, whether in whispering a syllable or in felling a forest” (Bynum, 1980). After processing sensory information from the environment, the motor system sends neural commands that cause coordinated, purposeful movements. These movements generally seem to occur automatically, outside our conscious attention. However, even simple actions consist of a complicated integration of brain’s several motor systems. The final purpose of this elaborate information processing is to enable us to interact with our environment.

It is generally believed that motor system is organised in a functional hierarchy. Imagine a familiar movement like typing on a keyboard. At the highest level that requires prefrontal cortex, the intention and desire to type is formed. This would be followed by a complex interaction with our current states and past memories to decide *what* word to type, *when* to start typing or *whether* to start typing at all (Haggard, 2008). The next level, which is concerned with the formation of a motor plan, involves interaction between the premotor cortex and sensory information from the parietal areas, as well as input from basal ganglia. The information in the parietal cortex about the environment and position of the body in the space serve to precisely adjust the internal representation of the motor plan in the premotor cortex. This information from action preparation processes are then converges on the primary motor cortex (M1) to execute motor commands by transmitting them to the spinal cord and to the muscles. The lowest level of the hierarchy thus involves spatiotemporal

details of the muscle contractions needed to execute the planned movement, which is passed thorough primary motor cortex as the 'final common path' (Sherrington, 1906). Interestingly, new evidence suggest that many of these processes can occur in parallel (Cisek, 2007). However, the ease with which we type words on a keyboard masks the true complexity of the processes involved. In normal conditions, we are not aware of the actual joint motion, muscle contractions and the action preparation processes. The true complexity becomes evident when we try to build machines than can perform human-like control of movement with the dexterity of a six-year-old child.

1.2.2. What makes an action voluntary?

Traditionally, movements have been divided into reflexes and those associated with voluntary actions. At the turn of the 19th century, Sherrington was the first to describe the reflexes as the basic units for movement. He proposed that reflexes are stereotypical movements that are generated in a linear, orderly fashion. Sensory inputs, by activating receptors in skin or muscle, automatically trigger a set of motor actions (output). For example, the most studies spinal reflex is the stretch reflexes: stretching a muscle activates receptors (muscle spindle), which in turn excites the muscles that contain it, through a spinal pathway.

In contrast to reflexes, volitional movements are the result of cognitive processes that enable the agent to exert some action on the outside world. The capacity of voluntary action control has enabled mammals to interact in sophisticated ways with their environment and to increase their chances in locating food, finding mate and avoiding predators. This capacity in humans, specifically, along with the fine anatomy of their hands and arms, has enabled them to change their environment and to create tools, giving them a distinct advantage over all other species.

Voluntary movements differ from reflexes in several important ways:

a) Voluntary actions are initiated by an internal decision to act and are not automatically triggered by external stimuli. In other words, the agent decides 'what' to do and 'when' to act internally, in the absence of any external cue to specify the action (Passingham et al., 2010a).
b) Voluntary actions are responsive to reasons: we often deliberate and consider reasons before performing a particular course of action (Anscombe, 2000). Even when we do not explicitly deliberate, we can give a reason to explain why we acted as we did.
c) The type, form and timing of voluntary actions are not directly determined by an external stimulus. They are thus 'free from immediacy' (Gold and Shadlen, 2007).
d) Voluntary actions often involve choices between alternatives, involving deciding whether to act or not. Thus, they have clear counterfactuals (Pereboom, 2011).
e) Once an intention to perform a voluntary action is formed it does not need to be executed immediately. A voluntary action could be prepared but hold to achieve some goal in the near or distant future. One can *control* whether, how and when to act.
f) Voluntary actions have context dependent association with sensory inputs. While specific external triggers always induce the same reflexive movement, same objects can invoke different voluntary responses depending on the context. Therefore, from a third person perspective, properties of a voluntary action are often unpredictable.
g) Compared to the fixed nature of reflexes, voluntary movements often improve with experience. We can always learn new skills and improve the ones we have already acquired.

While some voluntary movements such as self-initiated endogenous actions meet all the above criteria, some other, like pressing the brake pedal in response to a red light, show some properties of both voluntary and reflexive movements.

1.2.3. Volition and 'free will'

These mentioned properties of a voluntary action, altogether, generate a specific subjective experience: we feel that our voluntary actions are under our conscious control and that intention precedes action. This brings a dualistic view of endogenous causation. Neurobiology, on the other hand, treats actions as being the direct and determined consequences of specific states and processes in the nervous system, particularly in the frontal and prefrontal cortex of the brain. The problem of whether we are in control of how we act, and what this control involves, is what is called the 'free will' problem and is one of the oldest and hardest problems in philosophy and psychology. Most of us are natural *incompatibilists,* we believe that we do possess control over how we act and we assume that this is not compatible with our actions being causally determined in advance by factors out of our control (Pink, 2004). We naturally experience the strong subjective feeling that we could have chosen to act otherwise. But how can we be free to act in a *deterministic* system?

A system is said to be deterministic if given a full description of all its elements and laws at t_1, only one possible state can follow at t_2. This means that its future state is entirely defined by its initial conditions. The belief in causal determinism gained momentum in the 17th and 18th centuries with rapid advances in science and Newtonian physics. Since then the world has been pictured as a deterministic physical system which can be explained and predicted by mathematical equations. The generalization of this mechanistic system to human actions, however, has been challenged in at least two ways. Modern quantum physics has identified some events that are demonstrably indeterminate. Further, as regards determinism in human behaviour, neuroscientific studies with high-performance brain-computer interface have largely failed to identify the specific mechanisms causing human action.

On one view, these challenges are no threat to plausibility of causal determinism in human action for two reasons: Firstly, based on quantum theories the motion of sub-atomic particles are not predetermined and are to some degree probabilistic and chancy. However, chance

alone does not constitute freedom and it's all about randomness. Thus, it is not clear how a random motion can justify freedom in action. Secondly, we usually fail to distinguish between a *'deterministic'* and a *'chaotic'* system. A system is chaotic if it displays extreme sensitivity to initial conditions, a response popularly defined as 'butterfly effect'. A weather system is a classic example of a chaotic system. It is important to remember that a system can be both deterministic and chaotic: a system can completely obey physical rules but be impossible to predict either because we don't have access to all the data needed to do the calculation, or we have all the data and resources but the computations are difficult. Likewise, we have largely failed to exactly predict human behaviour, even though we are dealing with a deterministic system.

Classical incompatibilists believed that freedom over actions is not compatible with causal determinism and therefore we are free. However, a new form of incompatibilism is emerging, similarly claiming that freedom over actions is not compatible with causal determinism, while concluding that freedom is impossible. The 'free will' problem becomes even more significant when we look at the place of freedom in morality and society, in general.

1.2.4. Volition and society

One of the major consequences of believing in 'free will' is the idea of moral responsibility. We hold people accountable for their actions and we punish them for wrong doing. We feel regret and guilt when we do things we were not supposed to do. Interestingly, we don't have the same feeling for our feelings or desires (Pink, 2004). We don't usually feel guilty for having a feeling of hostility toward someone as long as we don't hurt them. This is because we don't feel in control of our feeling but we strongly feel in control of our actions. The idea of being a free agent lies at the heart of our moral thinking.

Importantly, 'voluntary action condition' underlies most systems of law. Standard common law hold one responsible for their actions if they perform the crime's 'guilty act' *(actus reus)* in combination with 'guilty mind' (*mens rea)*. However, for a person to be guilty, the *acteus*

reus must include a voluntary act. If I do a faulty act during an epileptic seizure I wouldn't be hold accountable for it. As required by law, an action is guilty if it is preceded by conscious intention to act. Unfortunately, even though these concepts are tightly linked with our understanding of the brain, very little is known about the brain basis of criminal responsibility (Maoz and Yaffe, 2014). As we will see in the next chapters, it has been shown that specific activity in the brain can be detected hundreds of millisecond before individuals become conscious of their intention to act (Libet et al., 1983). Another set of experiments showed that stimulating specific areas of the brain triggers an intention to move without actually causing the action (Desmurget et al., 2009). Even though these examples do not establish the inadequacy of the law, they suggest that neuroscience can inform our understanding of criminal responsibility.

1.2.5. Disorders of volition

The feeling of control over our actions is not only central to idea of moral responsibility but many neurological and psychological disorders involve abnormalities of volition. Here I will briefly review some the main neurological disorders of volition (Kranick and Hallett, 2013):

Tics are sudden repetitive movements and when multiple forms of them are present in patient they are called Tourette syndrome. Patients with tic/Tourette syndrome often feel their ticks are willed, but in response to an urge. However, patients often acknowledge that they are responsible for their actions. Interestingly, patients with strong urge before their tics show later conscious intention of their voluntary actions compared to a healthy control group (Ganos et al., 2015).

Lack of volition is hallmark of schizophrenic patients with delusion of control. These patients make normal movements but claim that their movement are not controlled by their own will but by an external agent. Compared to tic disorders, patients with delusion of control both deny willing the movement and responsibility for it. Computational models of motor control and sense of agency have been mainly successful in explaining positive symptoms of

schizophrenia (Fletcher and Frith, 2009; Voss et al., 2010). These models will be briefly described later in this chapter.

Alien hand syndrome is a rare neurological disorder in which patients often deny having control over the actions of one of their arms (usually the left arm). They report that the affected hand is 'alien' or is 'having a mind of its own'. Three main variants of AHS have been distinguished in the neuropsychological literature: frontal, callosal and posterior. The most common pathologies underlying AHS are corticobasal syndrome, stroke and Creutzfeldt-Jakob disease (Hassan and Josephs, 2016). This symptom will be important in Chapter 3.

Previous examples all included disorders with loss of volition. However, in some rare cases, patients have hyper volition. Anosognosia of hemiplegia is one of those rare cases that usually occur after a stroke in the right hemisphere. Patients are unable to move their arm due to hemiplegia, but once asked to try to move their arm, patients will insist that they are moving it and claim agency for the movement (Kranick and Hallett, 2013).

1.3. Brain circuits for voluntary action

1.3.1. Measuring volition

In this chapter I will review neural mechanisms underlying voluntary action control. Importantly, like any other scientific topic, studying volition in the laboratory requires appropriate tools for manipulating it, and methods for measuring it. However, unlike most experimental studies we cannot investigate the voluntary action system by giving it an input and measuring the output, because by definition voluntary actions are stimulus-independent (Haggard, 2008). Therefore, the classic techniques of system identification engineering cannot be readily applied.

Most previous studies on voluntary action consist of explicitly asking participants to perform an action (pressing a key) when they 'feel like it' or 'voluntarily', or to choose between several action alternatives at a specific time. These studies have been criticized for failing to capture volition and relying on paradoxical instructions to "be volitional in your action". This may encourage participants to demonstrate an artificial randomness in their behaviour (Jahanshahi and Dirnberger, 1999), simply because folk psychology associates volition with randomness. Recently, we developed a novel operational definition of voluntary actions by embedding endogenous "skip" responses that terminate unpredictable waiting for a perceptual decision. These are compared to a control block where participants are occasionally *instructed* to skip by an unpredictable visual cue (Chapter 2).

Despite limitations in methodology, recent neuroscientific evidence has suggested that human voluntary actions are special not just because of the experiences associated with it but because of their distinct neural mechanism.

1.3.2. One action system or two?

Classically, two cortical systems have been distinguished in driving actions (Passingham, 1987; Haggard, 2008; Passingham et al., 2010a). A parietal-premotor circuit that is involved in stimulus-driven actions and a fronto-median circuit that is responsible for voluntary actions. This *'volitional'* subsystem includes a circuit starting from multiple sites in the brain, gets filtered in basal ganglia, outputs to the frontal cortex, especially supplementary motor areas (SMA) and pre-supplementary motor areas (pre-SMA), and finally ending in primary motor cortex (M1). Passingham (1987) did one of the pioneering studies in this field by showing that monkey premotor areas direct actions based on information from external environment, while SMA directs actions based proprioceptive cues concerning animal's own actions. This was confirmed in another electrophysiological study in monkeys that found more cells firing in SMA compared to premotor cortex in self-initiated compared to externally-driven actions (Okano and Tanji, 1987). Interestingly, lesioning pre-SMA and SMA bilaterally

affected self-initiated movements in monkeys but not the externally-triggered movements (Thaler et al., 1995). Later papers on primate motor system challenged the idea of M1 as the 'final common pathway' and suggested that premotor areas are differentially involved in specific aspects of motor behaviour (Dum and Strick, 2002). While some are specialized in movement preparation, others are engaged in visually-cued movements.

A study by Deiber et al., (1999) was the first fMRI study in humans to dissociate these cortical circuits. Participants were asked to make self-initiated and visually-triggered movements in separate blocks. The authors observed extensive activity in frontomesial motor areas during self-initiated condition, including pre-SMA, rostral cingulate zone (RCZ) and the caudate cingulate zone (CCZ), suggesting functional specialization of these areas in self-initiated actions. Another neuroimaging study in humans with PET contrasted self-initiated with externally-cued movements (Jenkins et al., 2000). Similar to previous experiment, extensive activity was reported in SMA and cingulate cortex in the self-initiated condition. Further activation was found in dorsolateral prefrontal cortex (DLPFC). Interestingly, by using a similar paradigm, Cunnington et al., (2002) found no difference in overall activity in SMA and cingulate cortex between movement conditions. However, the timing of the hemodynamic response was significantly earlier in the pre-SMA for self-initiated compared to externally-triggered actions.

More evidence of a neurofunctional dissociation between intentional and stimulus-based actions comes from a study by Obhi and Haggard (2004). The authors measured electromyogram (EMG) during three keypress conditions: internally generated, externally-triggered and a truncation condition in which subjects were unpredictably instructed by a cue to make a keypress during an internally-generated condition. Reaction time (RT) to the external cue was significantly higher in the truncated compared to the externally-triggered condition. However, there was no effect of truncation on the EMG activity. They concluded that subjects were switching between two forms of actions preparation rather than modifying an ongoing action, suggesting an independent representation for internally and externally

guided actions. Functional dissociation between these two types of actions has also been reported in other behavioural experiments (Waszak et al., 2005; Welchman et al., 2010).

Recent evidence, however, suggests that the dichotomy between internally and externally-driven actions may not be as evident as previously thought, suggesting some level of interaction between the two. Nachev et al., (2008) argued that the reported functional dichotomy may be due to the more complex nature of self-initiated tasks. By nature, self-initiated actions contain conflict and therefore higher levels of complexity compared to externally-triggered actions, thus it is incorrect to assume a clear contrast between internally and externally-guided actions. This is incompatible with Passingham et al., (2010) who suggested that medial frontal cortex is crucially involved in self-generated action (see also Nachev and Husain, 2010; Passingham et al., 2010b). The distinction between voluntary and stimulus-driven actions was also questioned in a study by Hughes et al., (2011). Participants prepared a voluntary action with either left or right hand. While preparing the action, they were occasionally interrupted by stimulus requiring either left or right hand action. The authors found increased voluntary motor preparation in congruent trials where participants were preparing the same hand as the one required by the stimulus compared to incongruent trials. These findings show that the externally-driven system has access to voluntary system in medial frontal areas, suggesting an interaction between the two systems. These results are inconsistent with the findings of Obhi and Haggard, (2004) who showed that activation of the volitional system interferes with the stimulus-driven system.

Despite clear disagreements, there is a general consensus that SMA and pre-SMA have a key role in preparation of voluntary actions (Passingham, 1995; Cunnington et al., 2005; Krieghoff et al., 2011). Scalp recording in humans from this area has revealed a slow rising negativity that begins 1 s or more before the onset of endogenous actions (Kornhuber and Deecke, 1965; Deecke and Kornhuber, 1978). Since this ‘readiness potential’ does not occur before involuntary movements, it has been endorsed as the electro-physiological sign of planning, preparation, and initiation of voluntary actions (Kornhuber and Deecke, 1990). The

first strong evidence to suggest that this RP comes from SMA was an EEG source reconstruction study by Ball et al. (1999), which showed activity in mesial frontal areas up to 2 s before action onset. The role of SMA and pre-SMA in preparation and initiation of voluntary self-initiated movement and as a source of RP was further supported by an fMRI study during a self-initiated sequential finger to thumb movement (Cunnington et al., 2003).

1.3.3. The Bereitschaftspotential: a marker for voluntary action?

The Bereitschaftspotential or readiness potential (RP) was first described by Kornhuber and Deecke (1965). They recorded EEG signal while participants were asked to perform a spontaneous finger flexion at a time of their own choice. After averaging the recorded signal over many trials, time-locked to the moment of action onset, they observed an ongoing negativity staring around 1.5 - 2 seconds prior to the action onset. This readiness potential is maximal at the midline centro-parietal area and is derived from slow changing electrical brain potential with frequencies below 1 Hz. The origins, properties and the role of this signal in voluntary action generation have since been investigated in numerous studies (Shibasaki and Hallett, 2006; Jahanshahi, 2013).

One of the most famous of these studies is the so called Libet experiment (Libet et al., 1983). Similar to the original experiment of Kornhuber and Deecke, they asked participants to flex their finger at a time of their free choice. Additionally, participants were required to report the time when they first felt the urge or desire to move. Analysis of the readiness potential showed that it rises in the brain well before participants become conscious of their intention to move. Although this experiment has been widely criticized (Verbaarschot et al., 2015), its controversial results have since been cited as a pioneering evidence against the human capacity for free will (Bode et al., 2014).

The RP is distinguishable into two components: an early RP (RP1) that starts bilaterally and is best recorded at central electrodes like Cz. Scalp negativity increases sharply around 400 ms prior to action, making a boundary between the early and late components. This late

component (RP2) is more specific for the site of movement and is maximal over the contralateral central area. While RP1 reflects higher order factors such as movement preparation and intention, RP2 is more reflective of movement execution. In pathological conditions, RP1 is abnormal in Parkinson Disease and RP2 in hemiparesis (Shibasaki and Hallett, 2006). The specificity of the late RP in motor preparation has been studied as the lateralized readiness potential (LRP). LRP is measured by subtracting the potential recorded at C3 from that of C4 for both left and right hand actions separately. In one of those innovative studies, Haggard and Eimer (1999) investigated the causal relation between RP and action awareness. They found a positive relation between LRP latency and the perceived time of voluntary action (W time), suggesting that processes underlying LRP, but not RP, may cause our awareness of movement initiation. However, another study with larger sample size failed to replicate this finding (Schlegel et al., 2013).

Since then, numerous studies have investigated the causal origins of the conscious intention to move in the human brain and its relation with RP. Patients with damage in parietal cortex are unable to report the time when they become aware of their intention to move (Sirigu et al., 2004; see also Desmurget et al., 2009). Further, attention to the intention rather than movement increases activity in pre-SMA (Lau et al., 2004). Strikingly, in an fMRI experiment, the brain activity in these fronto-parietal areas were used to predict the outcome of a decision up to 10 s before subjects become aware of their decisions (Soon et al., 2008). In the only single neuron study of human volition, Fried et al., (2011) showed that a small group of neurons in the medial frontal cortex show a gradual ramp-like increase in firing rate before voluntary action that resembles RP in EEG. Interestingly, this increase in firing rate emerges 1.5 s before subjects become aware of their intention to move. These results are in general agreement with previous findings from EEG RP and previous reports that the experience of volition is based on readouts of brain activity in frontal areas. More recently it has been shown that amplitude of RP is correlated with haemodynamic response in the premotor

cortex (Zama and Shimada, 2015), and single trial analysis of RP-like signals has also been attempted (Schultze-Kraft et al., 2015).

Single trial analysis of RP is usually not feasible due to low signal-to-noise ratio in single-trial. Participants usually make several *free* actions, repetitively, choosing for themselves when to act each time. RP is then measured by averaging potentials over many trials. This method assumes that there is a single endogenous signal that gives rise to the classical ramp-like negativity of the mean RT but is covered by noise at single-trial level. Based on this interpretation, if one could control all external factors, RP could be recovered from single-trial EEG and would show a fixed onset and amplitude across all trials. This has been supported by a previous study by Dirnberger et al., (2008): Close to zero skewness showed that increase of negativity preceding a self-initiated action has a symmetrical distribution across trials. This suggests a fixed-onset model of RP, where the gradually rising negativity represents a fixed neural event prior to voluntary actions.

Having discussed the evidence in favour of classical interpretations of such RPs as biomarkers of a specific cognitive process, I now turn into new models showing that RPs could simply reflect stochastic fluctuations in neural activity. Two studies in particular, one using EEG in humans (Schurger et al., 2012) and the other using single-unit recording in rats (Murakami et al., 2014), involve a radical revision of the interpretation of the RP by showing that integration to bound models offer a plausible explanation for the gradual build-up shape of RP.

Evidence accumulation models have been traditionally used in perceptual decision making task. In these models sensory evidence and internal noise are integrated over time until a fixed threshold-level firing rate is reached, at which point the decision is made. Schurger et al., showed that the precise moment of an action onset can be explained by stochastic accumulator models. However, in the case of self-initiated voluntary actions, that are not externally cued, what drives the model is not sensory evidence but ongoing random

fluctuations in the neural activity. Hence, RP is simply a by-product of cross-trial averaging of autocorrelated noise. Precursor brain activity would not reflect goal-directed causes of action, but simply a by-product of constantly fluctuating neural activity. The same stochastic accumulator model was used to explain a spontaneous self-initiated action, in which rats decide when to abort waiting for delayed tone (Murakami et al., 2014). In line with this evidence, crayfish was also shown to demonstrate a gradual build-up activity prior to spontaneous transition to foraging (Kagaya and Takahata, 2010).

Consistent with this account Jo et al., (2013) showed that RP emerges through an unequal ratio of negative and positive slopes of slow cortical potentials prior to a voluntary action. They rejected that idea of RP as a neural marker of voluntary action preparation suggesting that it reflects spontaneous neural activity during the task. Further, it has been noted that technical details of how to analyse RPs are crucial to interpretation of the Libet-style experiments (Verbaarschot et al., 2015). For example, RP recordings are conventionally baseline-corrected using a baseline 2.5 until 2 s before action. This involves the implicit assumption that RPs begin only in the 2 s before action onset, but this assumption is rarely articulated explicitly, and is in fact questionable. I address this issue in more details in chapter 2. Recently, Alexander et al., (2015) revealed that RP is not uniquely related to motor activity and can be observed even in the absence of motor related processes. They suggested that RP may reflect a general anticipation effect rather than a specific marker of action preparation. In this regard, it has been shown that stimulus anticipation is preceded by a cortical negative wave, the contingent negative variation (CNV) (Walter et al., 1964). The CNV has been associated with expectation and temporal processing (Casini and Vidal, 2011; Van Rijn et al., 2011). Similarly, the stimulus preceding negativity (SPN) precedes an anticipated stimulus that provides task relevant information (Damen and Brunia, 1987; Brunia and Damen, 1988). Both CNV and SPN can occur for purely cognitive tasks in the absence of any concomitant movement.

Two theoretical papers have recently discussed the neural correlates of voluntary action in the light of the current research on RP. Moutard et al., (2015) distinguish two modes of cortical dynamics in the brain: a fast, high-amplitude ignition that is associated with stimulus-driven activations and an ultra-slow (<0.1 Hz), low-amplitude, resting-state activity. During sensory stimulation the system enters the active mode, where an increase in sensory parameters is associated with a nonlinear increase in neuronal response. The system passes a threshold; the signal is amplified and remains active for a while. In the absence of external stimuli, e.g., in the case of voluntary action, ultra-slow stochastic fluctuations accumulate gradually up to the threshold level and may spontaneously cross the threshold at a random time initiating a rapid signal amplification and action execution. Interestingly, based on this model, participants only become aware of their intention to act once the signal passes the ignition threshold.

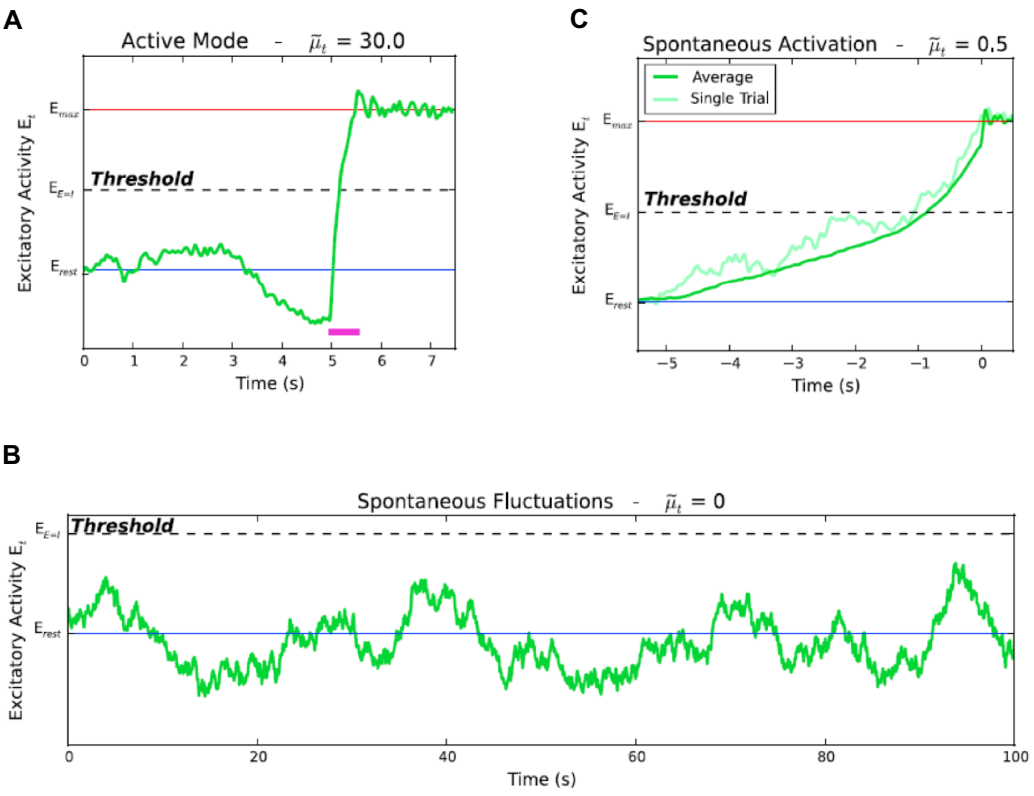


Figure 1.1.A. In response to external stimuli system enters the active mode. B. In the absence of sensory stimuli the system remains in resting-state demonstrating ultra-slow random fluctuations. C. These slow fluctuations may gradually accumulate and spontaneously cross the threshold. Adapted from Moutard et al., (2015).

In 'Action-related slow cortical potential sampling hypothesis', Schmidt et al., (2016) argue that RP reflects neither movement preparation, as suggested by the classical accounts of RP, nor decision processes, as suggested by Schurger et al. Instead they suggest that slow cortical potential can only increase/decrease the probability to initiate a voluntary action. Importantly, as evident from single trials that are preceded by a positive RP, these slow fluctuations do not integrate and do not need to cross a threshold for an action to happen. These fluctuations only *mediate* action initiation by giving rise to an *urge* to move, which itself is not action specific. The *urge* increases one's likelihood to initiate an action, but is not a necessary condition. Based on this argument one may wonder why self-initiated actions happen during some negative fluctuation but not others. They complete their hypothesis by suggesting that the *necessary condition* for voluntary action is not cortical potentials with negative fluctuations or an urge to move but one's *conscious decision*. However, they do not explain what this conscious decision is if not a readout of cortical activity preceding a voluntary action.

As discussed there is emerging evidence that electrophysiological signals prior to voluntary actions simply reflect ebb and flow of background neuronal *noise*, rather than outcome of a specific neural event. But what do we mean by neural noise and how can we differentiate it from desired signals?

1.3.4. Noise in the motor system

Noise is defined in Oxford English Dictionary as "random or unpredictable fluctuations and disturbances that are not part of a signal". Although noise affects every level of the nervous system, its role in neuronal computation and variability in behaviour has been widely ignored. There are several sources of noise in the nervous system (Faisal et al., 2008). The most

widely studies source of noise is *sensory* noise. Any noise at this level can be amplified by sensory receptors and increase the overall level of noise in the system. Therefore, animals spend a lot of energy to reduce noise at this level. At the next level there is *cellular* noise. There are different sources of noise at cellular level such as ion channels, synaptic vesicles and binding of signalling molecules to receptors (Faisal et al., 2008). Interestingly, different type of neurons in different regions show different levels of variability (Shadlen and Newsome, 1998). Finally, at the effector level there is *motor* noise, which is derived from noise in motor neurons and muscles.

Even though these sources of noise are mainly at microscopic level, they can affect the system at macroscopic level and even at the behavioural level. Therefore, the nervous system uses different strategies to minimise the negative consequences of noise in the nervous system such as averaging and prior knowledge (Faisal et al., 2008). However, it is not desirable to completely cancel out noise as it has been shown that optimal levels of noise can have positive consequences. For example, while at low level of noise the sensory signal may not reach the threshold, at higher optimal levels, noise helps a threshold-crossing event by increasing the overall level of activity in the system (Shu et al., 2003).

Importantly, regulating the level of noise and holding it at optimal levels is very costly for the central nervous system, both from a metabolic point and also the amount of time required to complete a task (Faisal et al., 2008). For example, it has been shown that performance accuracy in a task is inversely related with the average duration of response, because fast responses are more noisy (Fitts, 1954). Interestingly, performing a voluntary action in a spontaneous manner is usually not an easy task. Participants often find it easier and less demanding to act in response to an external cue compared to ‘free’, ‘self-initiated’ actions. One speculation is that to perform a voluntary action, in the absence of external stimuli, motor system spends lots of energy to control the optimal level of noise in the system, making self-initiated actions metabolically costly.

After all, what is noise and what is signal and how can we differentiate them? Maybe what we call noise just represents our lack of knowledge of the system (Drongelen, 2006). Human brain is a very complicated physiological system, it is a chaotic system, but is also deterministic. Therefore, all the random processes in the brain may in fact be specific signals, if we could discover the entire range of complex interactions that govern the system. In the next section I will discuss how some inspiring experiments have taken advantage of the variability in the system as part of the signal.

1.3.5. Variance as a signature of neural computation

Early studies in sensory neuroscience emphasised 'rate coding' as the principle of information coding in sensory systems. In this form of coding what matters is the total number of spikes over a period of time but not inter-spike intervals. Variability in the inter-spike intervals was thought as 'neural noise'. However, we now know that this variability in spike timing carries important information that is often referred to as 'temporal coding', suggesting that this variability, in fact, is an important part of the signal (Stein et al., 2005).

The first study to use a measurement of variability as an internal signal in the motor system is the study by Churchland et al., (2006). Monkeys were trained to reach to a target to receive reward. However, after introduction of the target, they had to delay their response until presented with a go cue. Measuring inter-trial variability of neural response in dorsal premotor cortex showed that variability was initially high but then converged with target onset. Variability reached its minimum after go-cue, just before action initiation. The authors suggested that brain optimises action preparation by bringing firing rates to their optimal value. Based on this 'optimal subspace theory', action preparation is complete when inter-trial variability declines to an optimal value, suggesting a consistent firing rate. This hypothesis was further supported by finding that RTs to reach a target was shortest when variability was lowest. Importantly, they found no relation between decline in variability and mean firing rate: for an action to happen, the mean firing rate does not need to rise to cross a certain threshold, it just needs to reach an optimal sub-space. This is in contrast with rise to

threshold models that predict shorter RTs with higher firing rates (Churchland et al., 2006; Shenoy et al., 2011). Interestingly, this decline in variability is not specific to premotor cortex. It is a widespread cortical effect that can be observed in different areas of the brain and in response to different types of stimulus, suggesting that cortical circuits become more stable by an input (Churchland et al., 2010).

More recently, it has been shown that this variability can be modelled as 'doubly stochastic'. *Doubly stochastic* models assume that inter-trial variability originates from variability in spike timing, which itself originates from variability in neural firing rate (Churchland et al., 2011). This doubly stochastic system was recently simulated by Litwin-Kumar and Doiron (2012) by using a deterministic model. Interestingly, even though their proposed model is deterministic, when placed in the network it exhibits stochastic firing rate and spiking. Hence it may be better to call their system *doubly chaotic*, because it is truly deterministic but seems stochastic because of the unpredictability of its output (Churchland and Abbott, 2012). Litwin-Kumar and Doiron's experiment nicely shows how a deterministic model could be used to simulate a stochastic process, highlighting the importance of across trial variability in complex network computations.

Other studies have investigated how neural variability could be influenced by internal physiological factors such as perception and memory (Marcos et al., 2013), or change of motor intention (Saberi-Moghadam et al., 2016). In the later study, monkeys were trained on a target-reaching RT task. Occasionally, while monkeys were reaching for target, its spatial location was changed and they had to update their intention and motor trajectory. Replicating previous findings (Churchland et al., 2006), across-trial neural variability in premotor cortex decreased after presentation of the first target, reflecting action preparation. This variability increased, and then decreased again, when the original motor intention suddenly changed. The magnitude of this change was correlated with the amount of time taken to correct the movement. This is the first study to probe inter-trial variability to decode changes of motor intention.

Analysing trial-to-trial variability to extract information from the system has also been attempted in human fMRI (He, 2013) and EEG/MEG data (Schurger et al., 2015). Overall, the mentioned studies in this section offer the compelling possibility that inter-trial variability provides an additional dimension to information coding in the brain. However, in all mentioned studies, the decline in neural variability is *triggered* when subjects are presented with a 'target'. In chapter 2, we will show that this modulation of variability can also occur in the absence of external stimuli, in the case of voluntary self-initiated actions.

1.4. Experiences associated with voluntary action

1.4.1. Intention

Our voluntary actions are often associated with two distinct experiences: the experience of *intention* and the experience of *agency*. The experience of intention reflects the initiation and control of the voluntary action, and possibly a prediction of the outcome. This conscious intention to act seems to be the cause of voluntary actions: most of us believe in a dualistic view that we do an action just because we intend to do so (Ryle and Dennett, 2000). Despite extensive research on precursors of voluntary action itself, which I previously discussed, the neural basis of conscious intention to act is less clear. However, one common view among neuroscientists is that both actions and intentions may be caused by a common neural system in the brain. We only assume that actions are caused by intentions because they are temporally correlated: actions are always preceded by intentions (Wegner, 2003a). More radical views suggest that conscious intentions are mere confabulations. After performing a voluntary action and receiving its outcome, we retrospectively infer the experience of intention to justify the cause of our action (Dennett and Kinsbourne, 1992). Despite being a hot topic of research, the neural correlate of intention is out of the scope of this thesis.

1.4.2. Sense of agency

Sense of agency is a fundamental aspect of human experience. It gives us the feeling that we are in control of our actions and can make changes in the world. It is fundamental to instrumental and goal-directed actions, and forms the cornerstone of humans' astonishing capacity to change their physical and social environment. Buildings, cities, technology, and almost all man-made structures would presumably be impossible without a sense of agency (Haggard and Eitam, 2015). Moreover, this feeling gives us a sense of responsibility and enables us to hold people accountable for their behaviours (Frith, 2014).. In the next few sections I will discuss this unique human capacity in more details and will later demonstrate how this experience, like most other human experiences, is a product of neural circuits in the brain.

1.5. What is the function of agency?

1.5.1. Sense of agency and the legal system

Sense of agency is not only fundamental to instrumental, goal-directed actions at the individual level, but also forms a cornerstone of everyday social life. In social animals, an instrumental action of one agent will have outcomes that affect other agents, so some social management of individual action is required. In particular, the sociolegal concept that each individual is responsible for their actions, and must answer for them to others, requires that individuals are first aware of their actions, and of the outcomes of those actions (Moretto et al., 2011; Frith, 2014).

Moreover, the legal system punishes agents based on how much they were in control of their actions at the time of wrongdoing, and therefore how much responsibility they had over those actions. By using an economic game, Singer et al., (2006) showed that participants did not reward or punish those players who were just following instructions even though they behaved like the players who acted feely. More recently, in a new version of Milgram's

experiment, Caspar et al., (2016) showed that when participants were ordered to perform a harmful action to their co-participant, they felt less in control of the consequences of their action compared to when they freely decided to do the harmful act.

1.5.2. Abnormalities of sense of agency

The importance of sense of agency is not limited to social contexts. Many neurological and psychiatric disorders involve abnormalities of agency (Kranick and Hallett, 2013). For example, patients with psychogenic hyperkinetic movement disorder (PMD) often experience a movement without associated sense of agency. More commonly, schizophrenic patients often complain of a lack of agency, and report the feeling that they are being controlled by an external force (Hauser et al., 2011; Moore and Fletcher, 2012). Alien hand syndrome is another rare neurological disorder with abnormalities of agency: patients often complain of lack of control of the actions of their hand, reporting that it has a 'mind of its own'(Della Sala et al., 1991; Hassan and Josephs, 2016).

1.6. What is the content of agency?

1.6.1. The ideomotor theory and inferential account of agency

Several theories have been developed to explain the content of sense of agency. In *ideomotor theories,* repeated association of actions and outcomes means that, over time, actions come to be represented primarily in terms of their anticipated outcomes or goal-states. By the same association, activation of the neural code for the goal event is then able to generate the voluntary action (Prinz, 1997). Based on this view the intermediate mental processes between action and outcome, such as action preparation and selection are no longer part of intention-action-outcome chain: it is enough to anticipate the effects of an action, for that action to be automatically triggered. Based on this account, we experience agency towards an action when our intentions matches with external outcomes. Therefore, internal volitional signals that precede action have no influence on agency experience:

agency is a post-hoc construct that is retrospectively inferred after the action is done (Metcalfe and Greene, 2007; Chambon and Haggard, 2013).

Wegner and Wheatley, (1999) showed that by priming participants just before an action it is possible to induce a false sense of agency. In his 'apparent mental causation' theory, Wegner suggests that people merely *infer* their own agency based on observing the combination of anticipatory thoughts and action outcomes given three conditions: thoughts precede the outcome, are consistent with them, and there should be no other obvious cause for the outcome. Therefore, there is no direct mental access to the internal processes that cause our actions, and the experiences of will and agency are mere inferences, or even illusions (Wegner, 2003b). Further empirical findings have since supported this theory (Wegner et al., 2004; Pronin et al., 2006).

1.6.2. Comparator models of agency

One of the most prominent models of agency is the so called 'comparator model' of the sense of agency. This model was originally developed as a theory of motor learning and motor control (Wolpert et al., 1995; Wolpert and Ghahramani, 2000), but has been increasingly used to explain the origins of agency (Frith et al., 2000; Blakemore et al., 2002). Compared to ideomotor theory, comparator model places great emphasis on intermediate processes of action generation. Based on this model, the internal forward model fist generates an efference copy of a motor command predicting the sensory consequences of the future action. This efference copy would be stored in the motor system and would be later compared with the actual sensory consequences of the action. If the predicted sensory information matches the actual sensory outcome, the action and the resultant outcome is recognized as self-generated. However, if there is a mismatch, sense of agency is lost and we would perceive another agent as the cause of action. In this model, agency is still generated retrospectively, after the action outcome. But compared to the ideomotor theory, it has explicit intermediate mechanisms and has been successfully used in numerous studies

to explain sense of agency in health and disease (Shergill et al., 2003; Synofzik et al., 2006; Voss et al., 2006) (Figure 1.2).

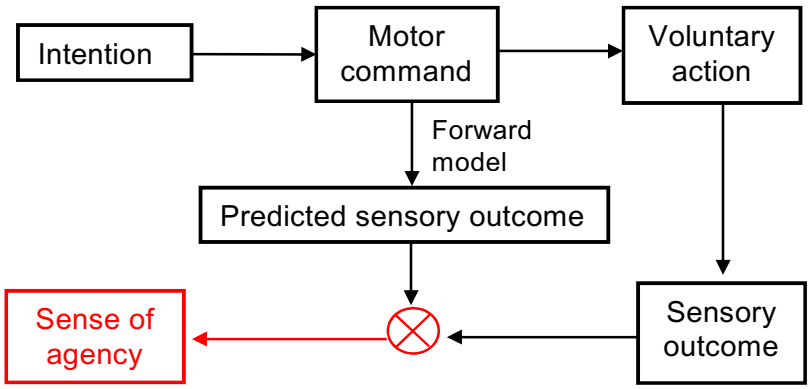


Figure 1.2. Comparator model of the sense of agency. An internal forward model uses efference copies of the motor command to predict outcomes. Sense of agency arises when there is a match between the predicted and actual sensory outcome of the generated action.

1.6.3. Prospective account of agency

Based on the comparator model of agency, sense of agency can only be formed when predictions are compared with reafferent feedbacks. Hence, agency is merely a retrospective phenomenon, happening only after an action is executed. However, in our daily life we feel in control of our actions even before receiving their sensory consequences. According to this 'prospective' account of agency, sense of agency can be generated prospectively, during the action selection process and before the outcome of action is known (Chambon et al., 2014b). This has been demonstrated in recent evidence that showed sense of agency could be modulated by subliminal priming. Importantly, this subliminal priming influenced agency during action selection process, without effecting the predictability of action outcomes (Wenke et al., 2010; Chambon and Haggard, 2012).

More novel frameworks suggest that sense of agency is influenced by both internal motor mechanisms and external factors. In a new approach, Moore et al., (2009b) changed the content of conscious thought prior to an action by supraliminal priming. They showed that priming modulates the perceived interval between voluntary and involuntary actions and their outcome. Interestingly, this effect was more prominent for involuntary actions, when no genuine intrinsic motor signal exists. They suggested that agency originates from a combination of internal cues, external cues and prior beliefs. Based on this 'Bayesian Cue Integration' approach, both internal sensorimotor signals and external contextual cues contribute to agency. However, the level of this integration depends on the reliability of each of these factors (Moore and Fletcher, 2012).

1.6.4. Active inference account of agency

Probably the most novel principle which has attempted to explain sense of agency is the 'active inference theory' (Wolpe and Rowe, 2014). According to this theory, at each level of cortical hierarchy there are *prediction* units which encode the *belief* at that level. The belief from each level is then sent to *prediction error* units at one level below. These prediction units compare the belief from levels above with sensory information at that level and calculate the *prediction error*, which is the difference between the predicted and actual sensory outcome. This error is then projects back to prediction units so they can adjust their prediction (Friston, 2010, 2011). According to this account, sense of agency arises when predictions (beliefs) at higher levels of motor system (e.g., pre-SMA) are consistent with sensory data at lower levels (e.g., M1) (Wolpe and Rowe, 2014). For example, abnormal awareness of action in psychosis or psychotic movement disorders may be due to disruption in normal balance between prior beliefs and the incoming sensory data (Edwards et al., 2012; Adams et al., 2013b). Active inference theory is thus far the most mechanistic model of agency.

1.7. How sense of agency is measured?

1.7.1. Two levels of agency

Despite its importance, it is very hard to measure sense of agency in laboratory settings. The most common way to measure agency is to ask participants to make a manual action. Participants would then see a visual feedback which may either be compatible with their own action or not. They will then judge whether or not they caused the visual event. Many varieties of this paradigm have been designed and implemented to measure sense of agency agency (Fink et al., 1999; Sirigu et al., 1999; Farrer and Frith, 2002; Farrer et al., 2003a, 2003b, 2008; David et al., 2007).

Synofzik et al., (2008) suggested that sense of agency comprises two different levels: an implicit 'feeling' and an explicit 'judgement' of agency. Based on this view, the subconscious 'feeling' of agency is produced at an implicit level by low-level sensorimotor signals. This non-conceptual 'feeling' of agency is then further processed by propositional representations, such as intentions and social cues, giving rise to the explicit 'judgement' of agency. Both feeling and judgement of agency contribute to the overall sense of agency, based on the context. While the traditional agency attribution paradigms can capture the *judgement* of agency, they are less successful in measuring the more internal, sub-conceptual *feeling* of agency.

1.7.2. Implicit measures of agency

To solve this problem, Haggard et al. (2002b) came up with a new design to measure sense of agency at an implicit level. Participants were asked to watch a rotating clock hand on the screen and to make judgments about the time of their actions and their consequent effects. The experiment consisted of four blocks. In the baseline conditions, participants either made a voluntary action or heard a beep, in separate blocks, while they were watching the clock rotating. They were asked to judge the time of their action or the beep. In the operant conditions, participants were asked to make voluntary actions (key presses), but this time

their key press produced a beep after 250ms. They were asked, in separate blocks, to judge the time of their key press or the tone that followed their action. These data were then used to compare the perceptual time of action and outcome in the baseline and operant conditions. The authors observed that when participants' key presses produced a tone (operant condition), the perceptual times of action and outcome moved towards each other and were bounded in time. They called this phenomenon *'Intentional binding'*. Importantly, they repeated the same experiment, but this time participants' voluntary movements were replaced by TMS-induced twitches. They observed a reversed binding effect, such that the perceptual interval between the action (involuntary twitches) and the outcome increased. Finally, the authors suggested that the same brain mechanism that is responsible for action-outcome binding, may be essential for the normal sense of agency (Fig.1.3). Since then, numerous studies have replicated the intentional binding effect (Moore et al., 2009a; Obhi and Hall, 2011; Takahata et al., 2012; Wolpe et al., 2013; Yoshie and Haggard, 2013). This effect has been interpreted both as a consequence of association between an action and its effect (Moore et al., 2009a), or as a multisensory cue integration (Wolpe et al., 2013). Intentional binding has been my implicit measure of sense of agency in chapters 3-6. Interestingly, implicit and explicit measures of agency do not always correlate, suggesting that they may measure different aspects of sense of agency (Ebert and Wegner, 2010; Saito et al., 2015).

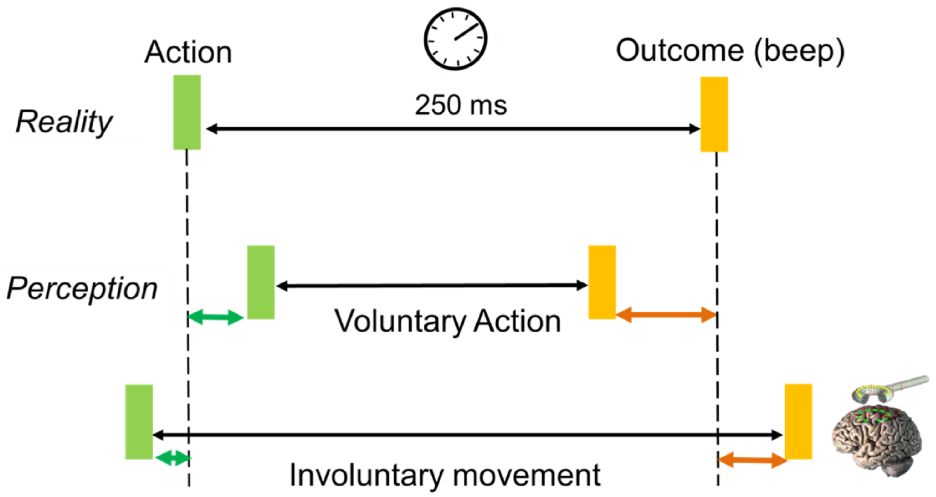


Figure 1.3. A schematic of intentional binding effect. The perceptual time of action and outcome moves toward each other for voluntary action, but not for involuntary movement.

1.8. What is the mechanism of agency?

1.8.1. Mechanism of explicit agency judgement

During the past few years many researchers have tried to shed light on the neural basis of voluntary action and two subjective experiences that accompany it, namely, awareness of voluntary action and the sense of agency. By studying patients with parietal and cerebellar damage, Sirigu et al., (2004) showed that in contrast to healthy volunteers and cerebellar patients, those with parietal damage report the time of their conscious intention to move (W judgement) very close to the actual time of movement. The authors suggested that a system in parietal cortex may be responsible for conscious monitoring of voluntary action. Based on this hypothesis, Desmurget et al., (2009) used direct electrical stimulation to evaluate the role of parietal areas in motor intention. They observed that stimulation of the parietal cortex induced a strong desire to act. When they increased the intensity of the stimulation, subjects reported performing the action. However, in reality, no movement had been performed based on the electromyographic data. This finding strengthened the claim that the activity in parietal cortex is responsible for conscious intention to act.

On the other hand, Fried et al., (2011) highlighted the role of frontal areas in willed action. Participants with implanted electrodes for epilepsy surgery were instructed to do the Libet's task (Libet et al., 1983) and to report the time of their 'urge' to move. Fried and colleagues observed an increase in firing rate in a group of medial frontal neurons, especially in supplementary motor area proper, before participants' conscious intention (urge) to act. The role of frontal and parietal areas in volition and their contribution to awareness of action is still a matter of debate (Haggard, 2011).

Several neuroimaging studies have explored volition in respect of the sense of agency. In one of the primary studies, Fink et al., (1999) used positron emission tomography (PET) to monitor cerebral blood flow while inducing mismatch between intention, proprioception and visual feedback of the healthy participants. The authors observed an increase in right dorsolateral prefrontal cortex (dLPFC) activity when participants had to control their hand while receiving incongruent visual feedback.

In another neuroimaging study, participants were asked to drive a circle along a path using a joystick. They were told that the circle would be either driven by themselves or by the experimenter. Higher activity was observed in anterior insula when participants attributed the movement to themselves, whereas external attribution of agency was associated with activity in inferior parietal cortex (IPC) (Farrer and Frith, 2002). In their next study, rather than contrasting self vs. other attribution, Farrer et al., (2003a) changed the degree of discrepancy between the movement and the visual feedback in four levels. In accordance with their previous findings, the activity in inferior parietal cortex increased with the degree of discrepancy and the subjects' feeling of being less in control of their actions. A reverse covariation was observed for the activity in insula.

To further investigate the role of parietal cortex in judgement of agency, Farrer et al., (2008) designed an fMRI study in which they manipulated the contiguity of action-outcome relationship and the authorship of the action, in separate experiments. For the first experiment, subjects were told to remove white golf pegs from a black grid with their hand, by looking at a screen which was projecting the image of the grid and the participant's hand, with a delay. After each movement, subjects reported if they perceived any delay in the outcome of their action. In the second experiment, subjects performed another set of movements and were told that the observed movements could be their own or another person's. After each movement, they judged whether the observed action was their own or not. The authors reported that the activity in right angular gyrus was associated with both awareness of action-outcome contiguity and the authorship of the action, suggesting that

right angular gyrus is responsible for conscious awareness of discrepancy between an intended action and its outcome.

By designing a novel virtual reality paradigm, Nahab et al., (2011) measured BOLD signal while modulating sense of agency during voluntary finger movements. As expected, participants lost their sense of control over their actions with an increase in feedback discrepancy. Based on the neuroimaging data, the authors identified two networks: a leading and a lagging network. They suggested that the leading network, including the right temporoparietal junction, is involved in mismatch detection. This information is then passed into the lagging network, including the bilateral prefrontal areas, which elevates the information from the leading network to conscious awareness.

In an attempt to dissociate the neural correlates of sense of agency and sense of ownership, Tsakiris et al., (2010) instructed participants to make active or passive hand movements and to view an image of their hand while resting inside fMRI. This image could be either synchronous or asynchronous with the hand movement. By analysing the neuroimaging data, Tsakiris and colleagues reported a dissociation between agency and body ownership. While activity in the midline cortical areas was associated with body ownership, activity in premotor areas, especially pre-SMA was associated with agency, supporting the independence of these two functions.

1.8.2. Mechanism of implicit agency: evidence from fMRI and brain stimulation

Most of the above mentioned neuroimaging studies were about explicit 'judgement' of agency, where agency was manipulated by inducing a state of incongruency between action and outcome. To investigate the implicit 'feeling of agency', Kühn et al., (2013) correlated the BOLD activity of the brain with the perceived action-tone interval (to replicate the intentional binding effect). They found a cluster in the left SMA proper, whose activity was correlated with the implicit measure of agency. However, they did not find any evidence in favour of

involvement of the angular gyrus. The authors concluded that frontal, and not parietal areas, are mostly involved with implicit measures of agency.

Although various neuroimaging studies have tried to illuminate the neural correlates of agency, there is a lack of scientific evidence for a causal association between brain activity and the sense of agency. By using continuous theta-burst brain stimulation (cTBS), Moore et al., (2010) investigated the neural correlates of intentional binding, as an estimate of sense of agency. They found that locally disrupting the activity of the pre-SMA, reduces the temporal binding of a voluntary action and its outcome, providing the first causal evidence of the contribution of medial frontal areas to the sense of agency. In another study, Chambon et al., (2014a) dissociated the prospective and retrospective components of agency using an established subliminal priming paradigm (Chambon et al., 2013) and investigated the causal role of left dLPFC and left IPC in the prospective component. Interestingly, disrupting IPC, but not dLPFC, at the time of action selection, disturbed the sense of control over subsequent outcome. The authors suggested that IPC contributes to both prospective and retrospective aspects of agency. This dual contribution makes IPC a suitable candidate for integrating various sources of information and cues.

Transcranial direct current stimulation (tDCS) is an emerging form of non-invasive brain stimulation that delivers weak current to the brain through a set of two electrodes. By inducing intracerebral direct current flow, tDCS is able to modify the neuronal excitability of the targeted brain region depending on the type of stimulation by a mechanism similar to long-term potentiation (LTP) and long-term depression (LTD) (Stagg and Nitsche, 2011). While anodal stimulation increases the excitability of the motor cortex, cathodal stimulation decreases it (Nitsche et al., 2008a; Reis and Fritsch, 2011). In comparison to TMS, tDCS is cheaper, more portable, and easier to apply. Importantly, its capacity to provide polarity specific brain modulation without causing action potentials, distinguishes it from other non-invasive brain stimulation methods (Filmer et al., 2014). Although early tDCS experiments were on the motor cortex, its modulating effects on behaviour and cognition are being more

widely recognized (Nitsche et al., 2008a; Jacobson et al., 2012b; Filmer et al., 2014; Douglas et al., 2015; Raja Beharelle et al., 2015; Hämmerer et al., 2016). In chapters 4 and 5 we will show how tDCS could be used to modulate brain circuits involved in control of action.

1.9. Aims of the current thesis

Brain activity prior to voluntary action has been a key focus of interest. However, mechanistic models of volition and sense of agency are lacking. The present thesis, therefore, aims to go beyond theoretical and conceptual account of volition and to come up with convincing neural mechanisms of voluntary action in human brain.

Chapter 2 develops a new approach to voluntary control based on endogenous decisions to opt-out of waiting for another task. We identify brain precursors of such "skip responses" to investigate *neural volitional signals*. We analyse *variability* of brain activity prior to repeated voluntary actions, rather than the traditional mean activity. This approach captures, for the first time, the key contrast between volition as a specific causal process in the brain, *vs.* volition as a random process reflecting "neural noise".

Chapter 3 investigates whether experience of voluntary action could be a reconstructive inference triggered by monitoring one's actions and their outcomes, or a read-out of brain processes related to action preparation, or some hybrid of these. We test this question in a group of healthy participants and in a patient with anarchic hand syndrome.

Chapters 4 and 5 aims to understand the mechanism of sense of agency by modulating brain circuits involved in control of action, while measuring stimulation-induced changes in intentional binding. We performed seven separate experiments using transcranial direct current stimulation, to investigate the contribution of frontal and parietal circuits in perceptual experience of agency.

In chapter 6, by pairing voluntary actions of one hand with involuntary movements of the other hand, we show that key aspects of agency experience can transfer from voluntary to involuntary movements. Our results explain why one can feel fully in control of one's actions even when they are performed automatically, without focal conscious attention.

Chapter 2

Neural events prior to self-initiated movements

A gradual buildup of electrical potential over motor areas precedes self-initiated movements. Classical interpretations of such 'readiness potentials' (RP) as biomarkers of a specific cognitive process of volition have been challenged by new models showing that RPs could simply reflect stochastic fluctuations in neural activity. Importantly, these models make contrasting predictions about the variability of neural activity during the premovement period. Participants waited for an unpredictably occurring dot-motion stimulus, and were rewarded for correct left-right manual responses to motion direction. They could alternatively decide to skip waiting, by making a bilateral keypress. The bilateral skip response thus reflects a purely endogenous decision to act, without any external trigger. In control blocks, the bilateral skip response was triggered by an unpredictable change in colour of the fixation point. We investigated whether scalp electrophysiological signals prior to action reflected a specific precursor process leading to voluntary action, or stochastic motor noise by analysing trial-to-trial variability of EEG prior to skip actions. Variability decreased prior to action in both conditions, because EEG epochs were action-locked. However, the decrease in variability became more marked for self-initiated compared to externally-triggered skip actions, beginning from 1.5 s prior to action. This convergence towards a fixed pattern suggests a preparatory process prior to self-initiated actions, while remaining entirely consistent with the possibility that preparation itself is triggered by random fluctuations. Indeed, computational modelling could account for the observed convergence of EEG by adding an additional linear build-up term to the original stochastic integrator model.

2.1. Introduction

Functional and neuroanatomical evidence has been used to distinguish between two broad classes of human actions: reactions to environmental cues, and self-initiated actions that happen endogenously, in the absence of any specific stimulus (Haggard, 2008; Passingham et al., 2010a). While there has been extensive research on action selection and initiation in response to external cues, the brain mechanisms for self-initiated actions are less often studied. Unlike externally-triggered actions, self-initiated actions are not time-locked to any external event. Most studies simply instruct human subjects to act 'freely', or 'spontaneously' (Libet et al., 1983; Jahanshahi et al., 1995; Haggard and Eimer, 1999; Cunnington et al., 2002). However, this instruction seems artificial, and could be interpreted in different ways.

We adapted for humans a paradigm previously used in animal research (Murakami et al., 2014), where endogenous actions are embedded in a wider landscape of reward-guided perceptual decision-making. Participants responded to the direction of unpredictably-occurring dot motion stimuli by pressing left or right arrow keys (Gold and Shadlen, 2007). Importantly, they could also choose to skip the trial, whenever they wished, by pressing both keys simultaneously. The skip response thus reflects a purely endogenous decision to act, without any direct external stimulus, and provides an operational definition of volition within this experimental design ((Murakami et al., 2014). Voluntary 'skip' responses were compared to a block where participants made the same bilateral 'skip' actions in response to an unpredictable visual cue.

The classical neural marker of precursor processes for endogenous action is the readiness potential (RP: (Kornhuber and Deecke, 1965). The RP is taken to be "the electro-physiological sign of planning, preparation, and initiation of volitional acts" (Kornhuber and Deecke, 1990) and was considered a pre-requisite of the conscious intention to act (Libet et al., 1983). The RP is typically calculated by averaging many trials, aligned to the moment of action – although single trial analysis of RP-like signals has also been attempted (Schultze-

Kraft et al., 2015). Most studies explicitly or implicitly assume that the RP reflects a putative 'internal volitional signal', with a constant, characteristic ramp-like form, which may be obscured by noise in any individual trial (Dirnberger et al., 2008). The RP would correspond to a specific process of preparation, having a fixed duration. This process would necessarily preceding action initiation. Indeed, end of the preparation process might be a causal trigger for action initiation.

However, the very idea that the RP reflects a specific precursor process leading to voluntary action has been recently challenged. Alternative models suggest that the early tail of the RP does not reflect a goal-directed process but rather reflects subthreshold stochastic fluctuations that influence the precise time of crossing the threshold for movement (Schurger et al., 2012; Murakami et al., 2014). Crucially, averaging these fluctuations with respect to action results in the appearance of a causal-looking buildup, like the mean RP. In fact, RP simply reflects cross-trial averaging of autocorrelated noise, rather than a goal-directed process that causes action. This view, which can be formally expressed in a quantitative model, and tested against neural data (Schurger et al., 2012; Murakami et al., 2014), involves a radical revision of the interpretation of the RP.

Controversies regarding precursor processes have been central to neuroscientific debates about volition (Libet et al., 1983; Dennett, 1985). We therefore asked whether electrophysiological signals prior to endogenous action reflect specific precursor processes for voluntary action, or are merely stochastic motor noise. These two models can both account for the morphology of the mean RP, but they offer different explanations of the distribution of premotor EEG across trials. On the stochastic model, neural activity eventually converges because of the combination of EEG temporal autocorrelation, and biased sampling time-locked to a threshold crossing event. On the classical model, the distribution of single trial RPs converges because the RP marks a specific precursor process that is directed at producing self-initiated action.

We found a prominent drop in inter-trial variability before self-initiated skip compared to externally-driven skip actions, suggesting that neural activity converges on a fixed pattern prior to self-initiated actions. Quantitative simulations using a stochastic integrator model showed that the distribution of EEG prior to endogenous action can be modelled as a combination of three processes: a contextual shift in EEG level, a stochastic random noise process, and a gradual linear buildup that emerges at least 1500 ms prior to action.

2.2. Materials and Methods

2.2.1. Participants.

24 healthy volunteers, aged 18-35 years of age (9 male, mean age = 23 years), were recruited from the Institute of Cognitive Neuroscience subject data pool. Two participants were excluded before data analysis (one could not finish the task and one left insufficient EEG data because of excessive blinking). All participants were right handed, had normal or corrected to normal vision, had no history or family history of seizure, epilepsy or any neurologic or psychiatric disorder. Participants affirmed that they had not participated in any brain stimulation experiment in the last 48 h, nor had consumed alcohol in the last 24 h. Participants were paid an institution-approved amount for participating in the experiment. Experimental design and procedure were approved by the UCL research ethics committee, and followed the principles of the Declaration of Helsinki.

2.2.2. Behavioural task and procedure.

Participants were placed in an electrically shielded chamber, 55 cm in front of a computer screen (60 Hz refresh rate). After signing the consent form, the experimental procedure was explained and the EEG cap was set up. The behavioural task was as follows: participants were instructed to look at a fixation cross in the middle of the screen. The colour of the fixation cross changed slowly and continuously throughout the trial. This colour always started from ‘black’ and then gradually changed to other colours in a randomised order. At the same time, participants waited for a display of random dots (displayed within a circular

aperture of 7º of diameter with a density of 14.28 dots/degree, initially moving with 0% coherence with a speed of 2º/s, (Desantis et al., 2016a, 2016b)) to move coherently (step change to 100% coherence) towards the left or right. They responded with the left or right hand by pressing a left or right arrow key on a keyboard, accordingly. The change in dot motion coherence happened abruptly. Correct responses were rewarded (2p). Conversely, participants lost money (-1p) for giving a wrong answer (responding with the left hand when dots were moving to right or vice versa), for responding before dots start moving, or not responding within 2 s after dot motion. The trial was interrupted while such error feedback was given. Importantly, the time of movement onset was drawn unpredictably from an exponential distribution (min = 2 s, max = 60 s, mean = 12 s), so waiting was sometimes extremely long. However, this wait could be avoided by a 'skip' response (see later). Participants could lose time by waiting, but get a big reward if they responded correctly, or could save time by 'skipping' but collecting smaller rewards (1p) (Fig. 2.1.A). The experiment was limited to one hour, so using the skip response required a general understanding of the trade-off between time and money. Participants were carefully informed in advance of the rewards for responses to dot motion, and for skip response, and were clearly informed that the experiment had a fixed duration of one hour.

There were two blocked conditions, which differed only in the origin of the skip response. In the '*self-initiated'* condition blocks, participants could skip waiting if they chose to, by pressing the left and right response keys simultaneously. The skip response saved time, but produced a smaller reward than a response to dot motion. A skip response thus reflects a purely endogenous decision to act, in the absence of any external instruction to act, and based on the tradeoff between later, larger, and smaller, earlier rewards. This provides an operational definition of volition within our experimental design, which captures some of the important features of voluntary control, as well as the linkage of voluntary action to other aspects of cognition, such as decision-making and judgement (Schüür and Haggard, 2011). Each block consisted of 10 trials. To ensure consistent visual attention, participants were

required to monitor the colour of the fixation cross, which cycled through an unpredictable sequence of colours. At the end of each block they were asked to give a rough estimate of the number of times the fixation cross turned 'yellow' (five alternatives: never, less than 50%, 50%, more than 50%, always). They lost money for giving a wrong answer. At the end of each block, participants received feedback of total reward values, total elapsed time, and number of skips. They could use this feedback to adjust their behaviour and maximise earnings, by regulating the number of endogenous 'skip' responses.

In the *'externally-triggered'* condition blocks, participants were exogenously instructed to make occasional skip responses *as soon as* the fixation cross changed colour to red (Fig. 2.1.B). The time of the 'red' colour appearance was yoked to the time of self-initiated skip responses in previous self-initiated blocks. The colour cycle of the fixation cross had a random sequence, so that the onset of a red fixation could not be predicted. The fixation cross ramped to 'red' from its previous colour in 300 ms. Again, a small reward was given for skipping. The trial finished and the participant lost money if s/he did not skip within 2.5 s from beginning of the ramping colour of the fixation cross. The 'red' colour was left out of the colour cycle in the self-initiated blocks. To control for any confounding effect of attending to the fixation cross, participants were also required to attend to the fixation cross in the self-initiated blocks and to report the number of times the fixation cross turned 'yellow' (see previous). Each externally-triggered block had 10 trials, and after each block feedback was displayed. Each self-initiated block was interleaved with an externally-triggered block, and the order of the blocks was counterbalanced between the participants. The behavioural task was designed in Psychophysics Toolbox Version 3 (Brainard, 1997).

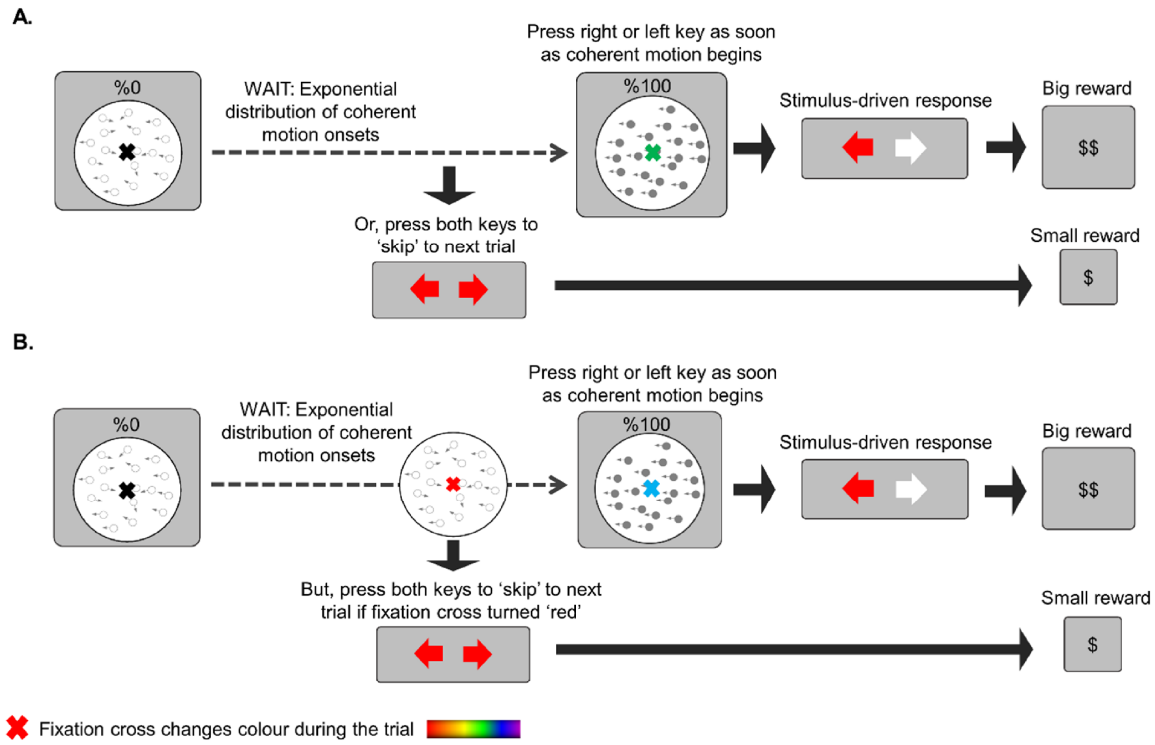


Figure 2.1. Timeline of an experimental trial. A. In the 'self-initiated' blocks participants waited for an unpredictably occurring dot-motion stimulus, and were rewarded for correct left-right responses to motion direction. They could autonomously decide to skip long waits for the motion stimulus, by making a bilateral keypress. They could lose time by waiting, but get a big reward, if responded correctly, or could save time by 'skipping' but collecting smaller rewards. The colour of the fixation cross changed during the trial. B. In the 'externally-triggered' blocks, an exogenous instruction to make a bilateral skip keypress was given by unpredictably changing the fixation point to 'red'. They got a small reward for skipping. Participants had to respond to dot motion direction, if the colour 'red' did not appear while waiting.

2.2.3. EEG recording.

While participants were performing the behavioural task in a shielded chamber, EEG signals were recorded and amplified using an ActiveTwo Biosemi system (BioSemi, Amsterdam, The Netherlands). Participants wore a 64-channel EEG cap. To shorten the preparation time we recorded from a subset of electrodes that mainly covers central and visual areas: F3, Fz, F4,

FC1, FCz, FC2, C3, C1, Cz, C2, C4, CP1, CPz, CP2, P3, Pz, P4, O1, Oz, O2. Bipolar channels placed on the outer canthi of each eye and below and above the right eye were used to record horizontal and vertical electro-oculogram, respectively. Impedance was kept below 10 kOhm during the experiment. EEG signals were recorded at a sampling rate of 2048 Hz and a bandpass filter of 0.01-100 Hz.

2.2.4. EEG preprocessing.

EEG data preprocessing was performed in Matlab (MathWorks, MA, USA) with the help of EEGLAB toolbox (Delorme and Makeig, 2004). Data were downsampled to 250 Hz and bandpass filtered to 0.01-30 Hz. All electrodes were referenced to the average of both mastoid electrodes. Separate data epochs of 4 s duration were extracted for self-initiated and externally-triggered skip actions. Data epochs started from 3 s before to 1 s after the action. To avoid EEG epochs overlapping each other any trial in which participants skipped earlier than 3 s from trial initiation was removed. On average, 5% and 4% of trials were removed from the self-initiated and externally-triggered conditions, respectively.

RP recordings are conventionally baseline-corrected using a baseline 2.5 until 2 s before action. This involves the implicit assumption that RPs begin only in the 2 s before action onset (Shibasaki and Hallett, 2006), but this assumption is rarely articulated explicitly, and is in fact questionable (Verbaarschot et al., 2015). We instead took a baseline from -5 ms +5 ms with respect to action onset. This choice avoids making any assumption about how or when the RP starts. Finally, to do artefact rejection, data epochs with values exceeding a threshold of ±150 μv were removed. On average 7% and 8% of trials were rejected from self-initiated and externally-triggered conditions, respectively. In the next step, Independent Component Analysis (ICA) was used to remove ocular artefacts from the data. Ocular ICA components were identified by visual inspection.

2.2.5. EEG analysis.

A typical RP-shaped negative-going slow component was observed in most participants. Preliminary inspection showed that this was generally maximal at FCz. Therefore, data from FCz was chosen for all analyses. Analysis was performed in Matlab (MathWorks) with the help of the FieldTrip toolbox (Oostenveld et al., 2010). We measured three dependent variables as precursors of both *self-initiated* and *externally-triggered* skip actions: mean RP amplitude across trials, variability of RP amplitudes across trials (measured by SD), and distribution asymmetry of RP amplitudes across trials (measured by skewness). To compare SD and skewness between the two conditions, data epochs were divided into four 500 ms windows, starting 2 s before action onset: [-2, -1.5 s], [-1.5, -1 s], [-1, -0.5 s], [-0.5, 0 s]. All p-values were Bonferroni corrected for four comparisons. To get a precise estimate of the standard error of the difference between conditions, paired samples t-tests were performed on jack-knifed data (Efron and Stein, 1981; Kiesel et al., 2008). Unlike the traditional methods, this technique compares variation of interest across subsets of the total sample rather than across individuals, by temporarily leaving each subject out of the calculation. In addition, we also performed cluster-based permutation tests on SD and skewness, which avoid some of the arbitrary assumptions associated with electrode and time-bin selection. The cluster-based tests were performed using the following parameters: time interval = [-2-0 s], minimum number of neighbouring electrodes required = 2, number of draws from the permutation distribution = 1000.

1.2.6. Modelling and simulations.

All simulations were done in Matlab (MathWorks). We used a modified version of the Leaky Stochastic Accumulator Model (Usher and McClelland, 2001), in which the activity of accumulators increases stochastically over time but is limited by leakage.

$$\delta x_i = (I - kx_i)\Delta t + c\,\xi i\sqrt{\Delta t_i}$$

Where I is drift rate, k is leak, ξ is Gaussian noise, c is noise scaling factor, and Δt is the discrete time step. This leaky stochastic accumulator has been used previously to model the neural decision of 'when' to move in a self-initiated task (Schurger et al., 2012). In that experiment, I was defined as the general urgency to respond (with a constant rate). This urgency, if appropriately small in magnitude, moves the baseline level of activity closer to the threshold, but not over it. Thus, urgency alone does not trigger action, but does increase the likelihood of a random threshold-crossing event triggering action. The threshold (β) was expressed as a percentile of the output amplitude over a set of 1000 simulated trials (50,000 time steps each).

In this experiment we used two models: The first model, (*Model 1*), is similar to that used by Schurger et al (2012). The parameters were fixed at the best fitting values previously used to model self-initiated actions (for details see (Schurger et al., 2012), Materials and Methods). We used c = 0.1, Δt = 0.001, k = 0.5, and I = 0.11. To make sure that these values also fit the current dataset, we separately found the best fit of the leak (k) and the drift rate (I) to the mean RP amplitude of each participant. The best fitting values for k and I were 0.5 (SD = 0.04) and 0.12 (SD = 0.02) respectively, very similar to the original model. Threshold (β) was obtained from the output of the first run of the model and was fixed at 0.44 (corresponding to the 99th percentile). '*Model 2'* has one additional parameter: the drift (I) now consists of two components: a step change I_1 (identical to I in Model 1: value 0.11), and a linear vector ($\boldsymbol{I}$) building from I_1 to I_2. I_2 is the only free parameter. All other parameters in Model 2 were fixed at the same values as in Model 1.

To perform the simulations, we followed the flowing steps: First, Model 2 was fitted against the real mean RP amplitude of each participant. A *least squares* approach was used to minimise root mean squared deviation (RMSD) by adjusting the best fitting I_2 of each participant (by using the MATLAB 'fminsearch' function). Second, Model 1 and Model 2 (by feeding it with the I_2 values found in the previous step) were run 22 times (equal to the number of participants) to simulate RP amplitudes in 1000 trials. Third, the standard

deviation (SD) across trials was calculated from the simulated RP amplitudes, for each model and for each participant. Fourth, the discrepancy between the SDs of the models' predictions and SDs of real data was quantified by RMSD. Fifth, the RMSDs of Model 1 against data was statistically compared with RMSDs of Model 2 against data with Wilcoxon signed-rank test in a 2 s time window before skip actions.

Importantly, this procedure fits the model to each participant's *mean* RP amplitude, but then tests the fit on the *standard deviation* (over 1000 simulated trials). The SD was obtained using the fitted parameters, and was compared to the actual measured standard deviation across individual trials of each participant's EEG traces. Fitting to means, and testing on SDs is a form of cross-validation. However, to establish the generalisability of Model 2, we additionally performed a 5-fold cross-validation. The data of each participant were randomly divided into calibration and validation samples. Model 2 was trained on the mean RP amplitudes of the calibration samples. Simulated SDs were obtained from the best fitting parameters of each calibration sample and were tested on the real SDs of the validation samples. Finally, the average RMSDs across validation samples of Model 1 was compared to the average RMSDs across validation samples of Model 2. This is an established method for comparing the performance of models with different number of parameters (Browne, 2000).

2.3. Results

2.3.1. Behavioural data.

On average participants skipped 108 (SD = 16) and 106 (SD = 17) times in the self-initiated and externally-triggered conditions, respectively. They responded to dot motion in the rest of the trials by pressing the left or right arrow keys (N = 177, SD = 61). The average waiting time before skipping in the self-initiated condition (7.3 s, SD = 1.6) was similar to that in the externally-triggered condition (7.6 s, SD = 1.6), confirming the success of our yoking procedure. The waiting time varied more across trials within each individual, than across

individuals, suggesting that volitional skip responses represented an on-line decision to act, rather than a pre-decided stereotyped response. Thus, the SD across trials of each individual participant had a mean of 3.17 s (SD across participants = 1.42 s) for self-initiated skips. Our yoking procedure ensured similar values for externally-triggered skips (mean of SD across trials 3.15 s, SD = 1.43 s). On average participants earned 214 p (SD = 33 p) from skipping and 278 p (SD = 99 p) from correctly responding to dot motion stimuli. This reward is in addition to the fixed amount they were paid for participating in the experiment. The mean and distribution of waiting time before skip action of each participant are presented in supplementary table 2.1 and Fig.2.2.

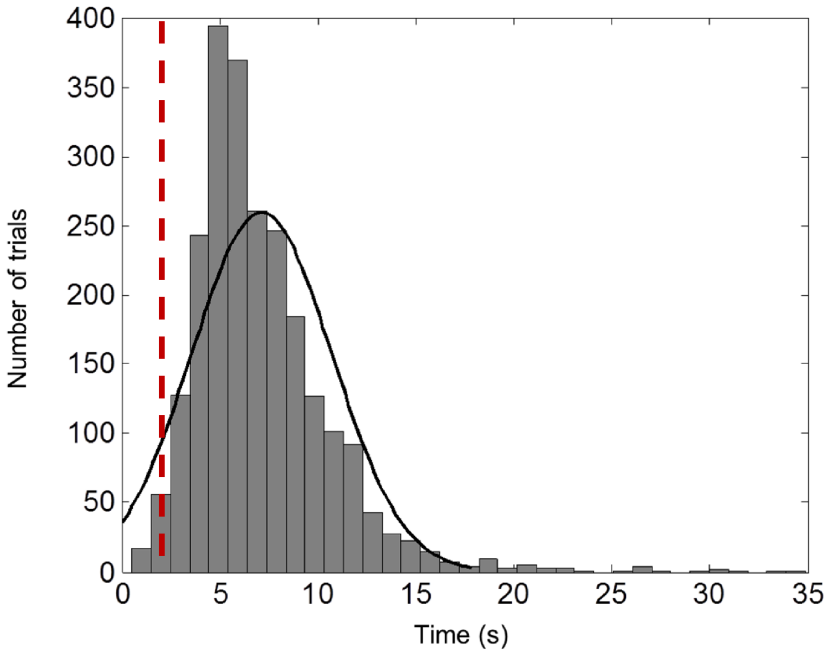


Figure 2.2 Time histogram of waiting time before skip actions in self-initiated condition across all trials and all participants. The dashed red line shows the optimum waiting time for getting the highest reward.

2.3.2. EEG variability decreases disproportionately prior to action in self-initiated and externally-triggered conditions.

EEG data from 22 participants were preprocesed and averaged across trials, separately for self-initiated and externally-triggered conditions. Figure 2.3.A shows the RP grand average

amplitude in both conditions. In line with previous evidence, the mean RP showed the negative-going shape for self-initiated actions. Note that our choice to baseline-correct at the time of the action means that the RP never in fact reaches negative voltage values. This negative-going potential is absent from externally-triggered skip actions (Papa et al., 1991; Jahanshahi et al., 1995).

The morphology of the mean RP might simply reflect the average of stochastic fluctuations, rather than a goal-directed build-up. However, these theories differ regarding the distribution of individual EEG trajectories across trials (see intro).

The standard deviation of EEG epochs decreased prior to skip action. This decrease is partly artefactual, since EEGs were time-locked and baseline-corrected to action onset. However, the premovement drop in EEG standard deviation was more marked for self-initiated than for externally-triggered skip actions. Paired sample t-test on jack-knifed data showed that this difference in SD was significant in the last three of the four pre-movement time bins before skip actions *(see materials and methods)*: that is from -1.5 to -1 s ($t(21) = 4.32$, $p < 0.01$, $d_z = 0.92$, p values are Bonferroni corrected for four comparisons), -1 to -0.5 s ($t(21) = 5.97$, $p < 0.01$, $d_z = 1.27$), and -0.5 to 0 s ($t(21) = 5.39$, $p < 0.01$,$d_z = 1.15$) (Fig. 2.3.B).

To mitigate any effects of arbitrary selection of electrodes or time-bins, we also performed cluster-based permutation tests *(see materials and methods)*. For the comparison between SDs prior to self-initiated vs externally-triggered skip actions, a significant cluster ($p < 0.01$) was identified extending from 1488 to 80 ms premovement (Fig. 2.3.C). This suggests that neural activity gradually converges towards an increasingly stable pattern prior to self-initiated actions. Importantly, this effect is not specific to FCz but could be observed over a wide cluster above central electrodes (Fig. 2.4). However, the bilateral skip response used here makes the dataset suboptimal for thoroughly exploring the fine spatial topography of these potentials, which we hope to address in future research.

Asymmetry of RP amplitude distribution was measured by skewness. While EEG values were positively skewed before skip actions in externally-triggered condition, their skewness remained close to zero for self-initiated skip actions (Fig. 2.5). However, the difference in skewness between the two conditions reached significance only in the 1 to 0.5 s premovement window ($t(21) = 2.90$, $p = 0.03$ corrected for 4 comparisons, $d_z = 0.62$), but not in the other three windows analysed. No significant cluster was detected by cluster-based permutation tests on skewness data.

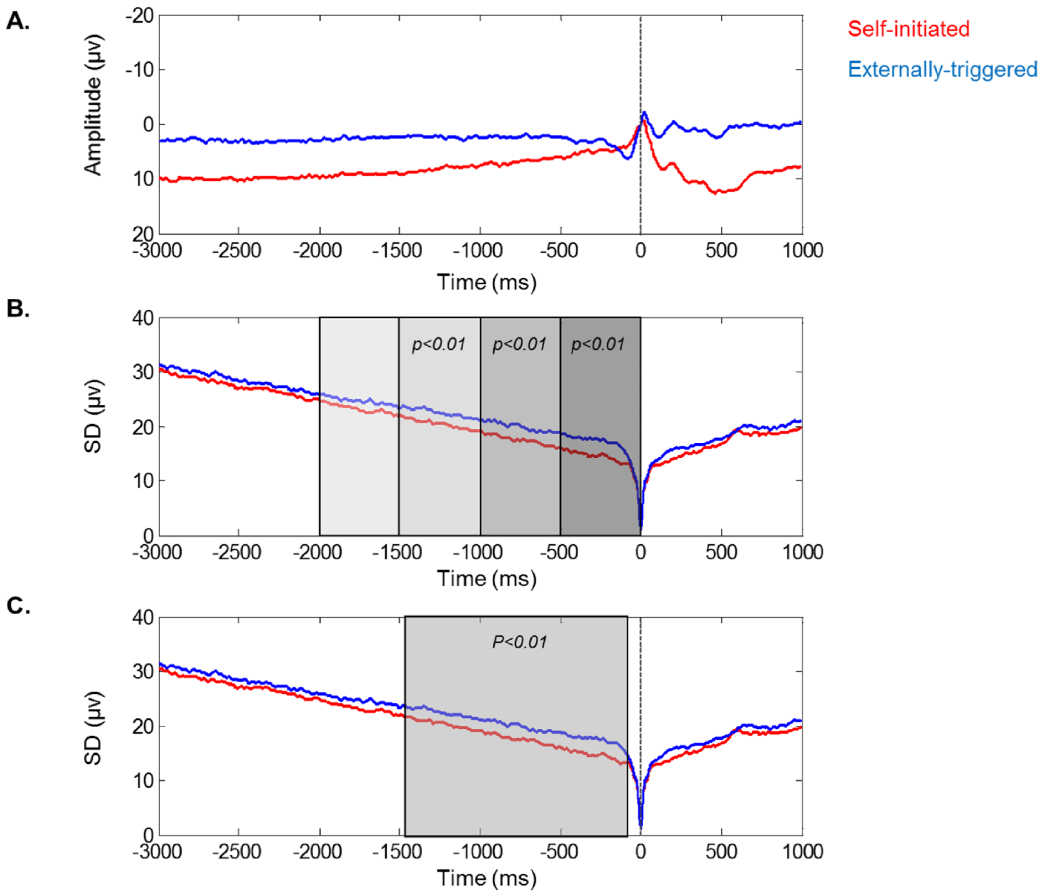


Figure 2.3. EEG activity prior to skip actions. A. Grand average RP amplitude. B. Standard deviation across trials averaged across participants (binned-based). Shaded areas show windows across which SDs were compared between the conditions. C. Standard deviation across trials averaged across participants (cluster-based). Shaded area show a significant cluster across central electrodes, detected by cluster-based permutation test. The red and blue lines represent self-initiated and externally-triggered skips, respectively. The dashed

line is the moment of skip action. Data are baselined to a 10 ms window around the skip and are recorded from FCz electrode.

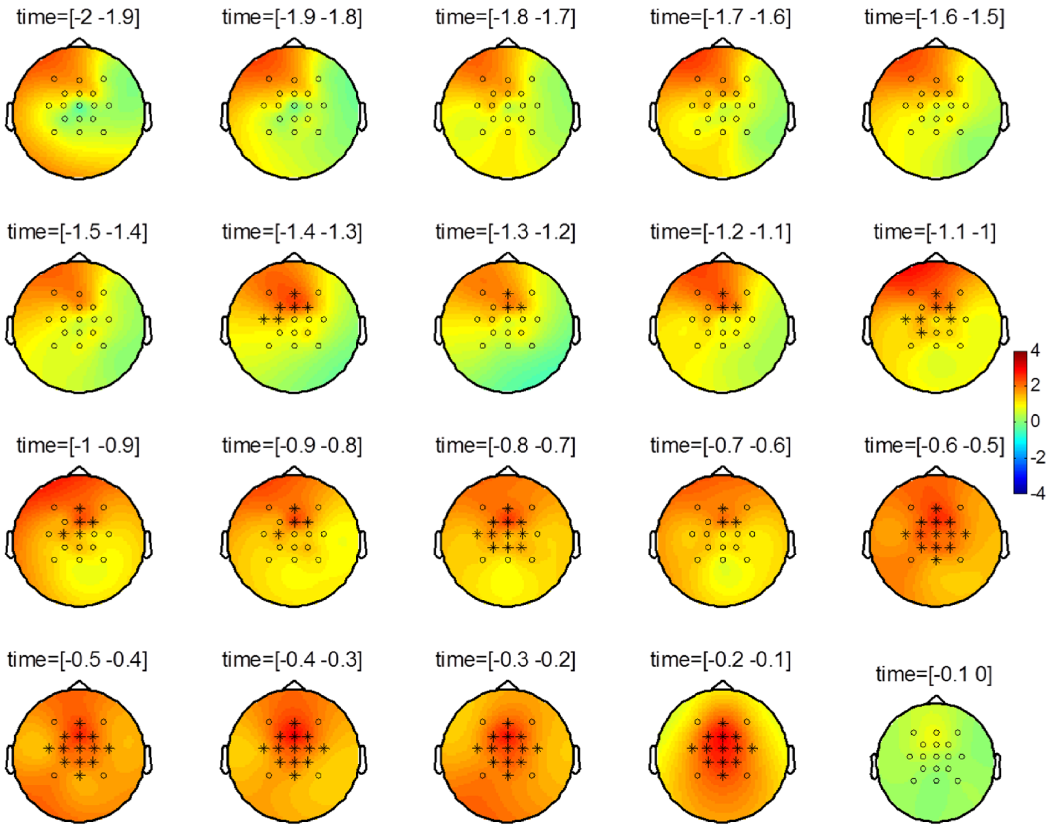


Figure 2.4. Topography of the difference in SDs between the two experimental conditions. The small circles represent EEG electrodes across which the permutation test was performed. The sensors which are members of a significant cluster have been marked*. Above each subplot its corresponding time interval is indicated.

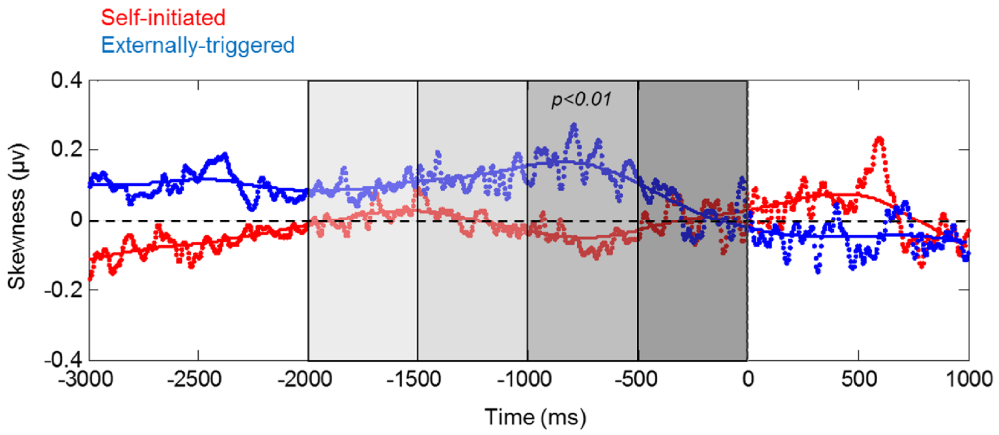


Figure 2.5. Distribution asymmetry of RP amplitudes across trials as measured by skewness. Shaded areas show windows across which skewness was compared between the conditions. The red and blue lines represent self-initiated and externally-triggered skips, respectively. The dashed line is the moment of skip action. Data are baselined to a 10 ms window around the skip and are recorded from FCz electrode.

It has been shown that stimulus anticipation is preceded by a cortical negative wave, the contingent negative variation (CNV) (Walter et al., 1964). The CNV has been associated with expectation and temporal processing (Casini and Vidal, 2011; Van Rijn et al., 2011). Hence, our measures in self-initiated condition could reflect both accumulating conditions that make a skip action desirable (e.g., passage of time without dot motion onset), and the preparation of the skip action itself. However, our externally-triggered skip condition controls for effects of mere passage of time, and expectation of dot motion onset.

To ensure that the key cognitive factors in the task were balanced between self-initiated and externally-triggered conditions, we also analysed mean and SD EEG amplitude prior to stimulus-triggered responses to dot motion (as opposed to skip responses). We did not observe any negative-going potential prior to dot motion (Fig. 2.6.A), again suggesting that temporal expectation did not strongly contribute to our ERPs. More importantly, the SD of EEG prior to dot motion onset did not differ between conditions in any time window ($p > 0.5$, Bonferroni corrected for four comparisons) (Fig. 2.6.B). This suggests that the disproportionate drop in SD prior to skip actions cannot be explained merely by a difference in expectation of dot stimuli or temporal processing.

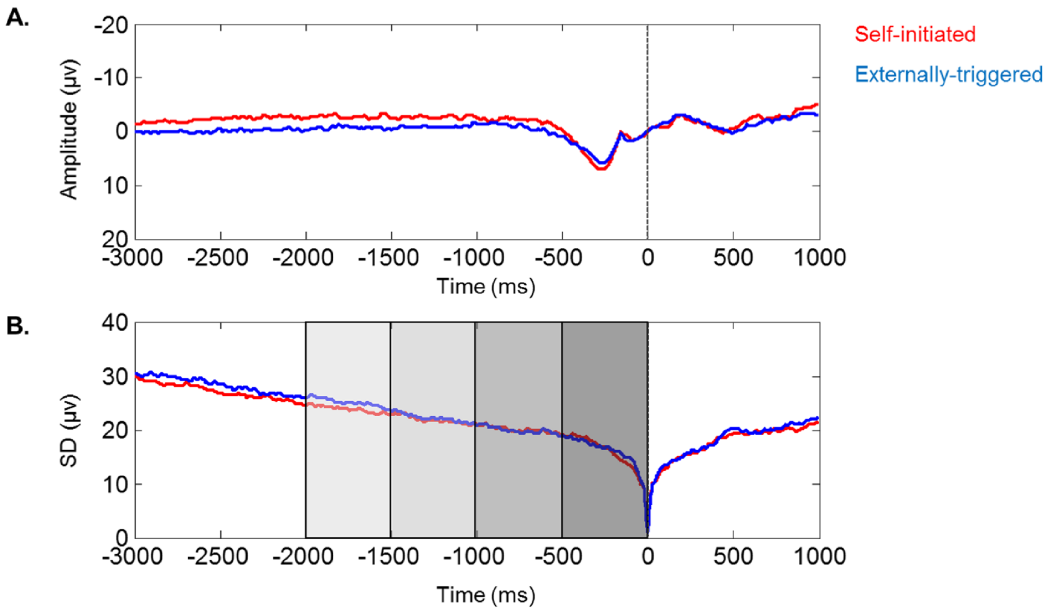


Figure 2.6. EEG activity prior to response to dot motion direction. A. Grand average RP amplitude. B. Standard deviation across trials averaged across participants (binned-based). Shaded areas show windows across which SDs were compared between the conditions. The red and blue lines represent activity in self-initiated and externally-triggered blocks, respectively. The dashed line is the moment of skip action. Data are baselined to a 10 ms window around the skip and are recorded from FCz electrode.

2.3.3. Modelling the converging EEG distribution of self-initiated actions.

The stochastic random fluctuation model is driven by internal physiological noise (Schurger et al., 2012). This model assumes that a constant primary drift shifts the premotor activity up closer to threshold and then a random threshold-crossing event provides the proximate cause of action (Model 1). Hence, the precise time of action onset is mostly random. However, the SD of cortical potentials in our dataset suggests that neural activity converges on a fixed pattern prior to self-initiated actions. We therefore modified the basic stochastic model (Model 1) by adding a preparatory process that might account for the fixed build-up to action. Specifically, we added a linear ramp progressing from the time of the primary step drift, up to the time of action onset (Model 2) *(for details see materials and methods)*. While

both Model 1 and Model 2 assume that self-initiated actions are preceded by random fluctuations, Model 2 assumes an additional gradual build-up of a fixed neural process. This build-up is summed with random fluctuations, and the combination of these leads to self-initiated action. To perform the simulations, the best fitting parameter was found by fitting Model 2 on the mean RP amplitude (Fig. 2.7.A and B). The best fitting values for I_2 were 0.73 (SD = 0.11) in self-initiated and 0.64 (SD = 0.19) in externally-triggered condition. In the next step, the expected SDs across trials were extracted from the simulated mean RP amplitudes of both models (Fig. 2.7.C and D). We then tested the goodness of fit of models by measuring the discrepancy between the observed SDs and the expected SDs under Model 1 and Model 2. The root-mean squared deviations of models from observed SDs were compared using a Wilcoxon signed-rank test. While Model 2 was as good as Model 1 in explaining the occurrence of externally-triggered skip actions (T = 175, p = 0.12, r = 0.33), Model 2 was significantly better than Model 1 in explaining self-initiated skip actions (T = 204, p = 0.01, r = 0.54).

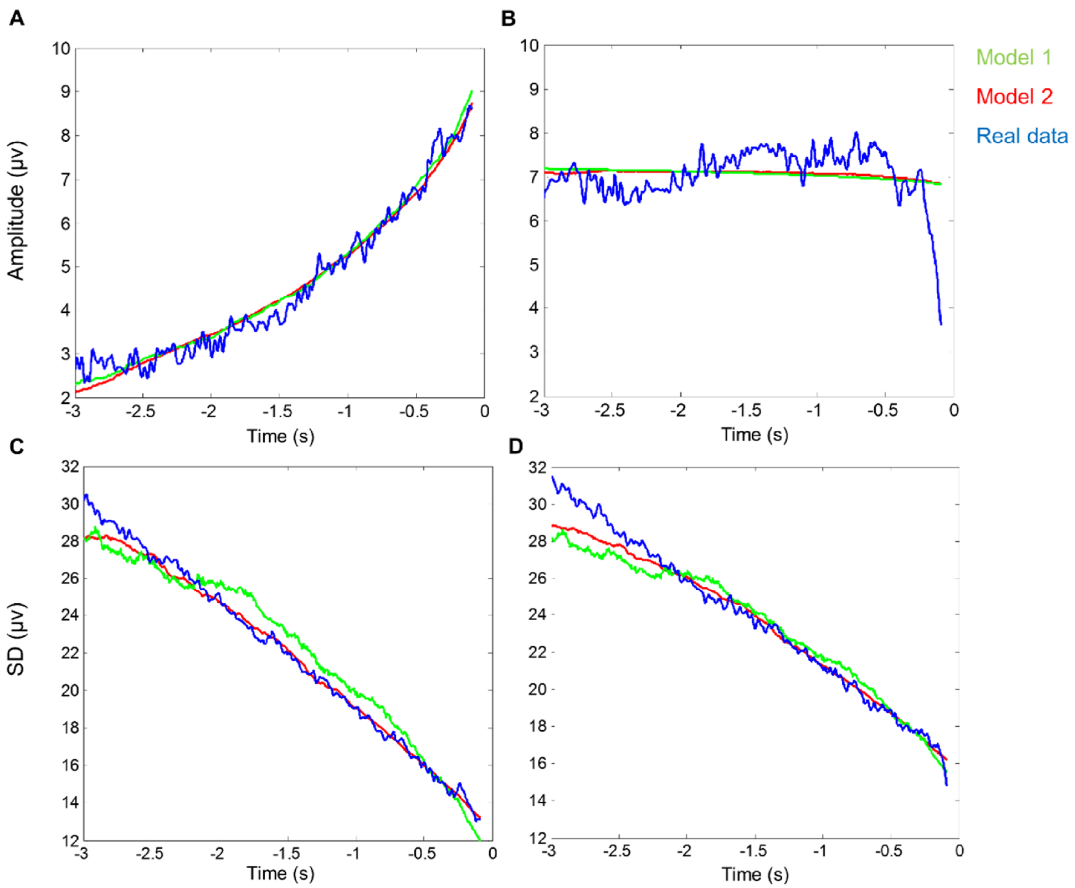


Figure 2.7. Simulated EEG data. A&B. Simulated grand average RPs from Model 1 and Model 2 plotted against the real data grand average in the 'self-initiated' (A) and 'externally-triggered' (B) conditions. C&D. Simulated SDs from Model 1 and Model 2 plotted against the real SD data in the 'self-initiated' (C) and 'externally-triggered' (D) conditions.

Finally, to establish the generalisability of Model 2, we performed a 5-fold cross-validation. The average RMSDs across test samples of Model 2 was significantly lower than average RMSDs across test samples of Model 1, for self-initiated skip actions ($T = 195$, $p = 0.026$, $r = 0.47$). This suggests that Model 2 can better predict the observed data even after controlling for its additional complexity (Fig. 2.8).

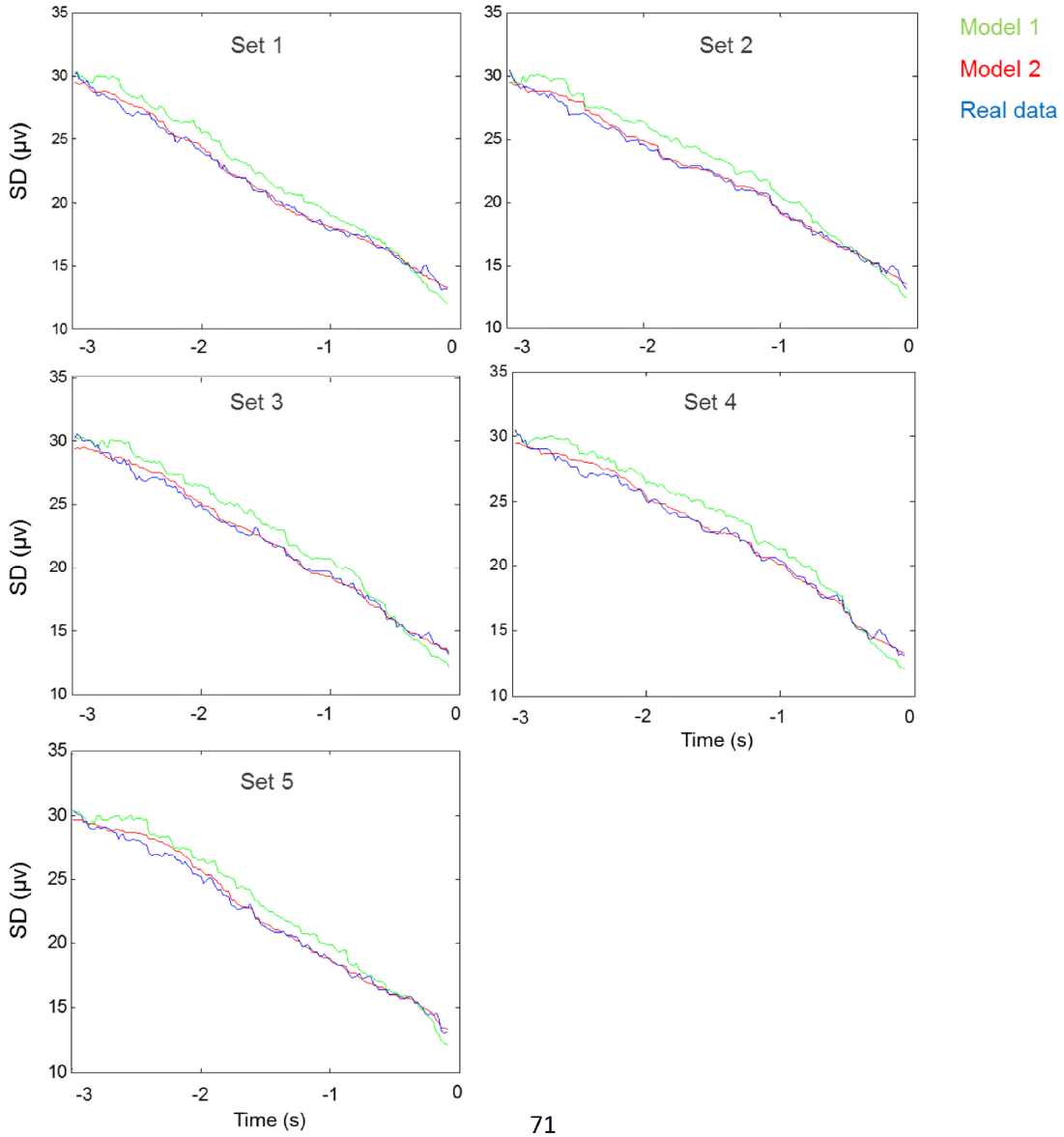


Figure 2.8. 5-fold cross-validation. Data of each participant in self-initiated condition was randomly divided into five calibration and validation samples. Models were trained on the mean RP amplitudes of the calibration samples and were tested on the real SDs of the validation samples.

2.4. Discussion

2.4.1. Skip when you are fed up waiting: operationalising volition as endogenous action.

The capacity for voluntary action lies at the heart of human nature, but the brain mechanisms that enable this capacity remain unclear. A key research bottleneck has been the lack of convincing experimental paradigms for studying volition. Many existing paradigms rely on paradoxical instructions to "be volitional in your action" (Libet et al., 1983; Haggard, 2005). In a novel paradigm, we operationalized voluntary actions as endogenous 'skip' responses embedded in a perceptual decision task, with a long, random foreperiod. Participants could decide to skip waiting for an imperative stimulus, by endogenously initiating a bilateral keypress. This action has features that have been considered hallmarks of volition, including internal-generation (Passingham et al., 2010a), reasons-responsiveness (Anscombe, 2000), freedom from immediacy (Shadlen and Gold, 2004), and a clear counterfactual alternative (Pereboom, 2011). Previous studies in animals also used *'Giving up waiting'* to study spontaneous action decisions (Murakami et al., 2014), but we believe this is the first use of a similar approach in humans. Crucially, operationalising voluntary action in this way avoids explicit instructions to "act freely", and avoids subjective reports. We compared such actions to an exogenous skip response triggered by a visual cue in control blocks.

2.4.2. Specific neural event or a background neural noise?

The neural activity that generates voluntary actions remains controversial. Several theories attribute a key role to medial frontal regions (Passingham, 1995; Nachev et al., 2008; Krieghoff et al., 2011). Averaged scalp EEG in humans revealed a rising negativity beginning 1 s or more before the onset of endogenous actions (Kornhuber and Deecke, 1965), and

appearing to arise in this area (Deecke and Kornhuber, 1978; Boschert et al., 1983). Since this 'readiness potential' does not occur before involuntary movements, it has been interpreted as the electro-physiological sign of planning, preparation, and initiation of self-initiated actions (Kornhuber and Deecke, 1990). RP-like brain activities preceding self-initiated actions were also reported at the single-neuron level (Fried et al., 2011).

However, this view has been challenged, because simply averaging random neural fluctuations that trigger a motor action also produces RP-like patterns (Schurger et al., 2012). Such stochastic accumulator models were subsequently used to predict rats' self-initiated actions in a task similar to ours (Murakami et al., 2014). Thus, it remains highly controversial whether the RP results from a specific precursor process that prepares voluntary actions, or from random intrinsic fluctuations. We combined an experimental design that provides a clear operational definition of volition, and an analysis of *distribution across individual trials* of pre-movement EEGs.

2.4.3. Neural activity converges towards a stable pattern prior to self-initiated actions.

EEG showed decreased trial-to-trial variability prior to skip actions. This partly reflects the time-locking and baseline-correction at the time of action: ERP methods necessarily imply zero variability at the baseline (Luck, 2005). However, around 1.5 s prior to skip actions, the decrease in variability became more marked for self-initiated compared to control externally-triggered skip actions. Since the skip action in the control condition has no volitional component, the decrease in variability prior to skip actions in the control condition presumably reflects only the effects of time-locking, and the temporal autocorrelation of the background EEG. However, the *additional* decrease in variability prior to self-initiated action may reflect convergence of neural activity towards a steady trajectory that precedes self-initiated actions. We hypothesised that this could indicate a specific preparatory process leading to voluntary action.

Measurement of variability has been extensively used in the analysis of neural data (Averbeck and Lee, 2003; Churchland et al., 2006, 2010; He, 2013; Schurger et al., 2015). Presenting a target stimulus decreases inter-trial variability of neural firing rate in premotor cortex (Churchland et al., 2006). Interestingly, RTs to external stimuli are shortest when variability is lowest, suggesting that a decrease in neural variability is a marker of motor preparation. Importantly, in previous studies, the decline in neural variability was *triggered* when monkeys were presented with a 'target', i.e. decreasing neural variability was triggered exogenously (Churchland et al., 2010). We have shown that inter-trial variability also decreases prior to a voluntary action, in the absence of any external target.

In our design, the onset and magnitude of gradual EEG convergence can be quantified by comparing the SD across trials prior to self-initiated versus visually-cued skip responses. These analyses characterise a precursor process prior to voluntary action. However, stochastic fluctuations may clearly also contribute to triggering action. We cannot directly estimate the magnitude of the stochastic contribution in our data, because measured across-trial SD depends on both stochastic fluctuations, and also on the inevitable statistical effects of time-locking and baseline-correction. These latter factors produce an artefactual appearance of convergence in any event-related potential, even for an entirely artificial 'event'. Therefore, our method cannot readily estimate the contribution of stochastic fluctuations to action decisions. In future research, we hope to directly compare the magnitudes of stochastic fluctuations and fixed precursor processes to voluntary action generation.

Integration to bound models have been recently used to account for the neural activity preceding self-initiated actions in humans (Schurger et al., 2012) and rodents (Murakami et al., 2014). Schurger et al.'s model first shifts premotor activity closer to a motor threshold. This is followed by a threshold-crossing event, triggered by stochastic fluctuation. The initial upward shift is operationalised as a step increase in activation level, defined in the model as 'a general urgency to respond'. In our experiment, however, general urgency to respond

should be similar between self-initiated and externally-triggered skip actions, since participants are in fact waiting for identical dot-motion stimuli in both conditions, and because the latency of skip responses that interrupt waiting was matched between conditions. Therefore, the decrease in trial-by-trial EEG variability, cannot simply be explained in terms of anticipation effects.

To account for the marked decrease in inter-trial variability we observed in self-initiated actions, we added an additional ramp-like build-up term to the original stochastic fluctuation model. While this enhanced model, which we called Model 2, was no better than the original model in explaining EEG patterns associated with externally-triggered actions, it was significantly better in explaining EEG associated with self-initiated actions, even after controlling for its additional parameters by cross-validation testing. The additional ramp-like buildup of Model 2, brings the activation level closer to the threshold. The ramp-like buildup in the model could correspond to a precursor process prior to voluntary action. It thus supplements, but does not abolish the role of stochastic fluctuation in triggering the action (Fig. 2.9).

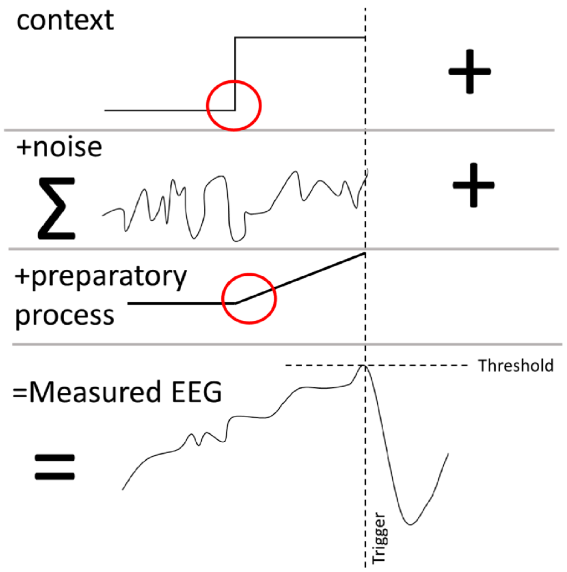


Figure 2.9. Quantitative simulations using a stochastic integrator model showed that the distribution of EEG prior to endogenous action can be modelled as a combination of three processes: a contextual shift in EEG level, a stochastic random noise process, and a gradual linear buildup of a preparatory process that emerges at least 1500 ms prior to action.

Clearly, any natural muscular action must have *some* precursors. Sherrington's final common path concept proposed that descending neural commands from primary motor cortex necessarily preceded voluntary action (Sherrington, 1906). However, it remains unclear *how long* before action such specific precursor process can be identified. Our result provides a new method for addressing this question. The question is theoretically important, because cognitive accounts of voluntary action control divide into two broad classes. In classical accounts, a fixed, and relatively long-lasting precursor process is caused by a prior decision to act (Kornhuber and Deecke, 1990; Anscombe, 2000). In other recent accounts, stochastic fluctuations dominate until a relatively late stage, and fixed precursor processes are confined to brief, motoric execution processes. Our study cannot show whether voluntary actions are caused by prior decisions, or by randomness. However, our results do suggest that the contribution of stochastic fluctuations is supplemented by a specific precursor process starting from around 1.5 s prior to action. Thus, our results suggest a novel, quantitative method for investigating the transitions between fluctuation and fixed process in the generation of voluntary action. Importantly, the precursor processes that our method identifies may be necessary for voluntary action, but may not be sufficient: identifying a precursor process prior to self-initiated movement says nothing about how often such a process is also present in the absence of movement. Hence this process might also be attributable to ongoing spontaneous fluctuations. For example, any putative last-minute decisions about whether to complete or veto the precursor process would be invisible in our experiment, because our design could not identify any vetoed precursor processes, or precursor-like processes that failed to result in a movement.

Our experimental design, and our EEG recording method allow us to place a lower bound on *when* precursors of voluntary action emerge, but not *where* in the brain. Importantly, the reduction in SD emerged around or before the time classically associated with the early RP1, rather than during the later periods classically associated with RP2 (Shibasaki et al., 1980), or with lateralised RP. Further, although we did not obtain any subjective reports in our

study, we note that the reduction in SD emerged substantially earlier than the time-point classically (and reliably) identified with subjective experience of volition (Libet et al., 1983).

Interestingly, our endogenous skip response resembles the decision to explore during foraging behaviour (Kolling et al., 2012; Constantino and Daw, 2015). That is, endogenous skip responses amounted to deciding to look out for dot-motion stimuli in forthcoming time-periods, rather than the present one. This prompts the theory that spontaneous transition from rest to foraging or vice-versa could be an early evolutionary antecedent of human volition.

Chapter 3

Changing the experience of a voluntary action

The experience of agency could be a reconstructive inference triggered by monitoring one's actions and their outcomes, or a read-out of brain processes related to action preparation, or some hybrid of these. Participants pressed a key with the right index finger at a time of their own choice, while viewing a rotating clock. Occasionally they received a mild shock on the same finger. They were instructed to press the key as quickly as possible if they felt a shock. On some trials, trains of subliminal shocks were also delivered, to investigate whether such subliminal cues could influence the initiation of voluntary actions, or the subjective experience of such actions. Participants' keypress were always followed by a tone 250ms later. At the end of each trial they reported the time of the keypress using the rotating clock display. Shifts in the perceived time of the action towards the following tone, compared to a baseline condition containing only a keypress but no tone, were taken as implicit measures of sense of agency. The subliminal shock train enhanced this "action binding" effect in healthy participants, relative to trials without such shocks. This difference could not be attributed to retrospective inference, since the perceptual events were identical in both conditions. Further, we tested the same paradigm in a patient with anarchic hand syndrome. Subliminal shocks again enhanced our measure of sense of agency in the unaffected hand, but had a reversed effect on the 'anarchic' hand. These findings suggest an interaction between internal volitional signals and external cues afforded by the external environment. Damage to the neural pathways that mediate interactions between internal states and the outside world may explain some of the clinical signs of anarchic hand syndrome.

3.1. Introduction

On one view, the experiences of volition and agency are post-hoc inferences, triggered by monitoring one's actions and their outcomes. In this case, preparatory brain events that precede action should not influence this experience (Wegner and Wheatley, 1999), although a "prior conscious thought" about acting may be necessary to trigger such inferences (Wegner, 2003b). Alternatively, experience of agency could depend on a readout of brain processes in frontal (Fried et al., 2011) and/or parietal areas (Desmurget et al., 2009) that precede voluntary action. Importantly, these two views make different predictions about how external stimuli might influence the experience of agency: If experience of agency is merely a reconstructive inference, interventions which influence brain processes preceding a voluntary action should have no influence on one's sense of agency, unless those interventions generate some perceptual event which can figure in the inference. On the other hand, if experience of agency depends on internal precursor signals that drive voluntary action, any intervention that influences these signals may also affect experience of agency, whether the intervention is consciously perceived or not.

In neuroscience, voluntary actions are often linked to a medial frontal pathway associated with internally-generated movement, as opposed to a parietal-lateral frontal pathway for reacting to external stimuli (Passingham et al., 2010a). Human experiments drawing on this tradition usually require participants to perform actions at a time of their own free choice, though this approach has been criticised for lack of ecological validity (Schüür and Haggard, 2011). Intervening on volition in such paradigms is methodologically difficult, because the experimenter cannot know when the participant will act. Further, any experimental intervention on precursor processes should preserve the 'internally-generated' aspect of voluntary action, rather than switching to a reactive mode of responding. Subliminal priming offers one potential method for studying volition. For example, subliminal visual primes have been used previously to manipulate the sense of agency by increasing the fluency of action

selection processes (Chambon and Haggard, 2012). Priming can "nudge" the brain towards selecting one action rather than another (Eimer and Schlaghecken, 1998). Compatible priming also increases sense of agency, as if the prime had made the action more strongly intentional. However, subliminal visual priming paradigms require a precise temporal relation between prime and a supraliminal 'go' signal. They therefore involve externally-triggered rather than internally-generated voluntary actions. Here, we used a novel design with subliminal electrocutaneous stimuli as a probe to influence brain processes preceding a voluntary action. We investigated how experimental manipulation of putative precursor signals can change the experience of agency in healthy adults and in an individual with 'anarchic hand syndrome' (AHS).

Healthy participants were asked to make voluntary key presses with their right index finger at a time of their own choosing. They occasionally received a mild electrocutaneous shocks on the same finger, and were instructed to press the key in reaction to such shocks as quickly as possible. This instruction aimed to set up a facilitatory association between shock and action. Both voluntary and reactive keypresses were followed by a beep 250ms later. Participants judged the time of the keypress using a rotating clock display. A shift in the perceived time of the action towards the following tone, compared to a baseline condition containing only a keypress but no tone, has been proposed as an implicit marker of agency (Haggard et al., 2002b). Crucially, the shift in action awareness towards the subsequent tone appears to reflect volitional signals, since it is absent for involuntary movements (Cravo et al., 2009), and increases with the amount of information that participants must generate internally (Barlas and Obhi, 2013).

Further, we delivered a train of subliminal shocks in some trials selected at random. We reasoned that the subliminal shocks might influence brain processes preceding voluntary action, because of the established association between shock and keypress. Because subliminal trials contained the same *perceptual* events as voluntary trials without subliminal shocks, any inferential processes should operate identically on both trial types. However, if

sense of agency depends on a readout of brain processes that precede voluntary actions, and if these processes can be influenced by subliminal stimuli, we might expect subliminal shock trains to affect sense of agency, as measured by intentional binding. Since the classical effect of subliminal priming is to facilitate voluntary actions, and since we included other supraliminal shock trials specifically involving such a link, we predicted stronger binding for voluntary actions on trials with subliminal shocks, compared to trials without subliminal shocks.

We also tested the same paradigm with a single patient, TP, with anarchic hand syndrome. AHS is a rare neurological disorder characterized by abnormal voluntary control over a limb (Kranick and Hallett, 2013). Three main variants of AHS have been distinguished in the neuropsychological literature: frontal, callosal and posterior. The most common pathologies underlying AHS is corticobasal syndrome, stroke and Creutzfeldt-Jakob disease (Hassan and Josephs, 2016). Patients often describe their affected arm as ‘alien’ or ‘having a mind of its own’. The movements are often goal-directed and triggered by external stimuli, but the patients are not able to control or stop them (Moore and Fletcher, 2012). Cognitive neuropsychologists have generally interpreted signs and symptoms of AHS using ‘object affordance theory’. Affordances are properties of objects in the environment which promote or invite action (Gibson, 1986; McBride et al., 2013). In healthy individuals, excessive reactivity to external stimuli is usually supressed by endogenous control mechanisms within medial frontal cortex (Sumner and Husain, 2008). Accordingly, impairment of these control mechanisms in AHS lead to patients becoming excessively responsive to external stimuli, even when they do not intend or wish to respond to them (Riddoch et al., 1998; McBride et al., 2013). In particular, the patients with AHS, including patient TP studied here, often involuntarily grasp external objects.

Cognitive neuropsychological accounts emphasise a form of “negative volition”, in which the lesioned cortex would have the normal role of ensuring tonic voluntary suppression of latent responses to environmental affordances. No studies, to our knowledge, have investigated

how the damage underlying AHS influences the processes that *generate* voluntary action itself. One hypothesis, consistent with neuronal (Fried et al., 2011) and areal (Filevich et al., 2012) evidence of intermingled action-promoting and action-suppressing representations in medial frontal cortex, predicts the damage that leads to AHS should also affect the generation, and experience of voluntary action.

3.2. Material and methods

3.2.1. Participants.

47 healthy volunteers, aged 18-35 years of age (14 males, mean age = 22.4 years, SD = 3.9), were recruited from the Institute of Cognitive Neuroscience subject data pool. All participants were right handed, had normal or corrected to normal vision, had no history or family history of seizure, epilepsy or any neurologic or psychiatric disorder. Participants confirmed that they had not participated in any brain stimulation experiment in the last 48 h, nor had consumed alcohol in the last 24 h. Participants were paid an institution-approved amount for participating in the experiment. Experimental design and procedure were approved by the UCL research ethics committee, and followed the principles of the Declaration of Helsinki.

TP is a 54 year old, right handed woman. She is a former secretary with 11 years of education. Twenty-three months before the testing session, she had a ruptured aneurysm of the right anterior cerebral artery, resulting in subarachnoid haemorrhage, involving the genu and trunk of the corpus callosum. After embolization, she had a vasospasm of the right middle cerebral artery. The most recent MRI (14 months before the testing session) showed damage in the corpus callosum (genu, body and splenium) and in the right anterior frontal and right basal frontal cortex, involving the anterior and middle cingulate gyrus (Fig. 3.1 & Table. 3.1).

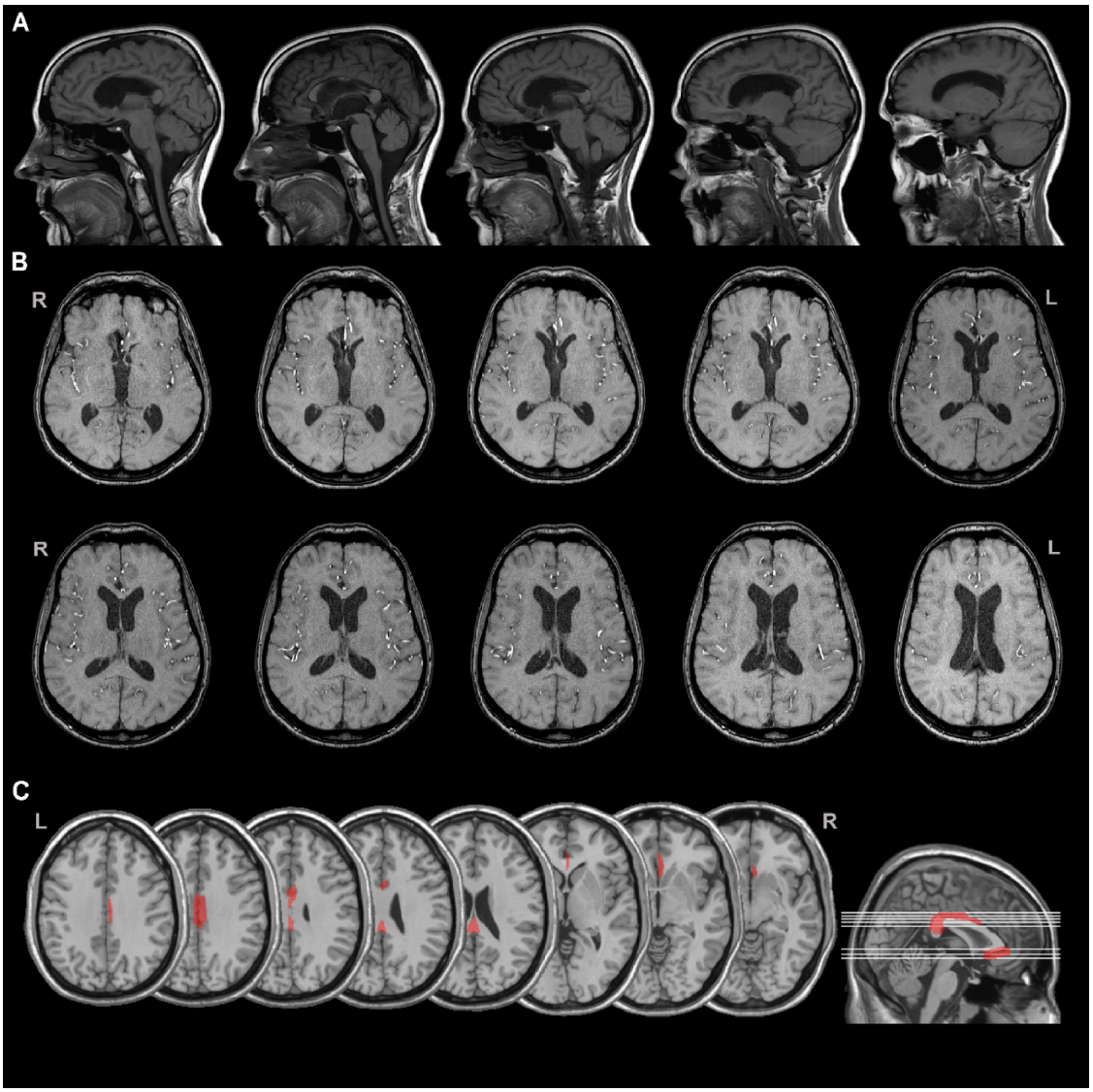


Figure 3.1. MRI scans of the patient TP in sagittal (A) and horizontal (B) view. C. Patient's lesion reconstruction. Mapping of the brain lesions was performed by MRIcro (Rorden and Brett, 2000). Lesions, as documented by the most recent MRI, were traced on the T1-weighted template MRI scan from the Montreal Neurological Institute provided with the MRIcro software.

Cortical Areas	**Voxels**	**Area %**
Olfactory_R	347	15%
Frontal_Mid_Orb_R	13	0%
Cingulum_Ant_L	86	1%
Cingulum_Ant_R	589	6%
Cingulum_Mid_L	234	2%
Cingulum_Mid_R	1050	6%
Cingulum_Post_L	58	2%
Cingulum_Post_R	35	1%
White Matter Areas	**Voxels**	**Area%**
Unclassified	2281	0%
Genu of corpus callosum	111	1%
Body of corpus callosum	664	5%
Splenium of corpus callosum	714	6%

Table 3.1. For each brain region, the number (first column) and the percentage (second column) of lesioned voxels are shown. Quantitative estimate of the damaged brain regions and white matter areas was performed by superimposing the traced lesion reconstruction on the "automated anatomical labelling" template (AAL) (Rorden and Brett, 2000), and on the John Hopkins University (JHU) white matter labels atlas (Rorden et al., 2007).

A complete neuropsychological examination at the time of the testing session showed residual attentional deficits (in the subtests "alertness", "optical vigilance", "acoustical vigilance" and "divided attention" of the Italian version of the Test of Attentional Performance: (Zimmerman and Fimm, 1992; Zoccolotti et al., 1994), mild impairments in perspective memory (Rivermead Behavioural Memory Test: (Wilson et al., 1985) and abstract classification abilities (Wisconsin Card Sorting Test: (Heaton et al., 2000). No additional impairments were found in working memory (sub-test "Working Memory" of the Italian version of the Test of Attentional Performance: (Zimmerman and Fimm, 1992; Zoccolotti et al., 1994)), long-term and short-term verbal memory (Buschke-Fuld Test: (Buschke and Fuld, 1974; Spinnler and Tognoni, 1987), Digit Span: (Orsini et al., 1987) and spatial

memory (Rey-Osterrieth Complex Figure Test: (Osterrieth, 1944; Caffarra et al., 2014), Corsi-Block tapping test: (Spinnler and Tognoni, 1987), or logical and reasoning abilities (Raven's Progressive Matrices: (Spinnler and Tognoni, 1987). In addition, TP showed no sign of apraxia with either limbs or hands (Test of limb apraxia: (De Renzi et al., 1968, 1980). She had mild paresis with hypotonia of her left upper limb and hand. Both superficial and deep sensitivity were normal. TP complained of her left hand behaving in an uncontrolled manner. For example, she reported that her left hand threw a towel into the bathtub full of water or that she blew her nose with a napkin that she held in her left hand, instead of using the handkerchief in her right hand. Despite the lack of voluntary control of her left hand, TP never denied ownership of the hand.

All experimental procedures were exactly the same in TP and healthy participants. The only difference is that all the data were collected from the right hand of the healthy participants, while for TP data was collected from both the right (healthy) hand and the left (affected) hand, in separate sessions.

3.2.2. Experimental procedure.

After filling the consent form, the general experimental procedure was explained for the participants. Non-painful electrocutaneous shocks were delivered from a programmable Digitimer DS5 Bipolar Constant Current Stimulator (Digitimer Ltd., Welwyn Garden City, UK). Cloth electrodes (Biosense Medical, Chelmsford, UK) were placed on the proximal and medial phalanx of the right index finger and were connected to the anode and cathode cables, respectively. Intensity of the shocks depended on the trial type (see later). The duration of each shock was set at 10 ms.

The behavioural task started after setting up the electrodes. Each experimental session consisted of three tasks: First, a detection task was used to detect the lowest threshold level at which participants were able to detect the shocks. Then, supra- and subliminal levels of shock were calculated from the threshold measure and a signal detection task was

administered to confirm perception of the shock stimuli. Participants who did not pass the signal detection task were excused and did not proceed to the next step. Finally, the intentional binding task was administered to provide a proxy measure of sense of agency (Moore and Obhi, 2012).

3.2.3. Threshold detection task.

An ascending staircase approach was used to detect the lowest levels at which participants were able to detect the shock (Moore et al., 2010). Shocks started at 0.1mA and increased in steps of 0.1 mA until the shock was detected, and then decreased in steps of 0.05 mA until the shock was missed, and then increased again in steps of 0.01 to find the detection threshold. A tone was played at the time of each shock and participants were asked to report if they felt a shock at the time of the tone or not. The level for supraliminal shock stimuli was set at 130% of the threshold level. The subliminal level was determined by reducing one step (0.01 mA) from the threshold (e.g., if the detection threshold was 0.45 mA, the subliminal level would be 0.44 mA). This strategy was chosen to ensure that subliminal shocks had sufficient energy to influence brain processes, while remaining imperceptible (see below).

3.2.4. Signal detection task.

The estimated supra- and subliminal shock levels were then validated in a signal detection task. Each signal detection task consisted of four types of trials in a randomised order: 20 subliminal shock trials, 20 subliminal catch trials (with no shock), 20 supraliminal shock trials and 20 supraliminal catch trials (with no shock). In each trial participants heard two tones, 5 s apart. They received a supraliminal shock at a random time between those two tones in supraliminal shock trials. No shock was delivered in catch trials. In subliminal trials, a train of subliminal shocks were delivered every 1 s starting from the first tone and ending with the second tone. At the end of each trial participants were asked to report if they felt any shock between the fist and the second tone or not. At the end of the task, participants' responses were used to estimates the sensitivity index (*d'*) for the supra- and subliminal shocks. To

proceed to the next step, participants were required to obtain a *d'* value within the range of 0.5-1.5 for subliminal shocks and a *d'* of >= 3 for the supraliminal shocks. If their *d'* did not match this criteria, the threshold detection task was repeated to find a new threshold followed by a signal detection task. If the desired *d'* was not achieved after four attempts, participant was excused and did not proceed to the intentional binding task.

3.2.5. Intentional binding task.

We used intentional binding paradigm as an implicit measure of agency. The task was based on previous studies (Haggard et al., 2002b), and was programmed in LabVIEW 2012 (Austin, Texas). Participants viewed a clock hand rotating on a computer screen, located 60cm in front of the participants in a quiet room. The initial clock position was random. Each full rotation lasted 2560 ms. Participants made voluntary keypress by pressing the enter key with their right index finger. Participants chose for themselves when to make the voluntary actions. After each key press, the clock hand stopped at a random location, participants made a time judgement according to condition (see later). Each experimental session consisted of two conditions, presented in separate blocks. At the beginning of each block, brief instructions for the relevant condition were displayed on the screen. In the *baseline* condition, participants had to press the enter key at a time of their own free choice. The clock hand stopped after 1500-2500ms (at random), and participants then judged the clock hand position at the time of their keypress. In this condition, participant's actions produced no sensory outcome and they received no shock. In the *agency* condition, participants were again asked to press the key at a time of their own free choice. However, this time each keypress produced a pure tone (1000 Hz, 100 ms duration) after 250 ms and they sometimes received a mild shock on their right index finger before pressing the key. At the end of each trial, participants made two subjective reports. First, they reported the clock hand position at the time of their keypress. Second, they reported whether they had felt a shock or not. Each block in the agency condition consisted of two types of trials in a randomised order: in two thirds of the trials a single supraliminal shock happened at a

random time, drawn from an exponential distribution (min = 1 s, max = 10 s, mean = 5 s) (Fig. 3.2.A, B). In the other one third, a 1 Hz train of subliminal shocks occurred starting from a random time within 500 ms from the beginning of the trial and continuing for 10 s (Fig. 3.2.C). The train ensures that any keypress occurs within 1 s of a shock. In all trials of the agency condition, participants were asked to press the enter key whenever they felt like but to press the key 'as quickly as possible' if they felt a shock. There were two possible outcomes in trials with a single supraliminal shock: either participants waited long enough, received the supraliminal shock and reacted (Fig. 3.2.B), or they voluntarily pressed the key before the occurrence of the supraliminal shock, in which case the supraliminal shock was cancelled (Fig. 3.2.A). The former trials were categorised as *'reactive'* trials, if participants accordingly reported feeling the shock, and the later trials were categorised as *'voluntary'* trials, if participants accordingly reported not feeling a shock.

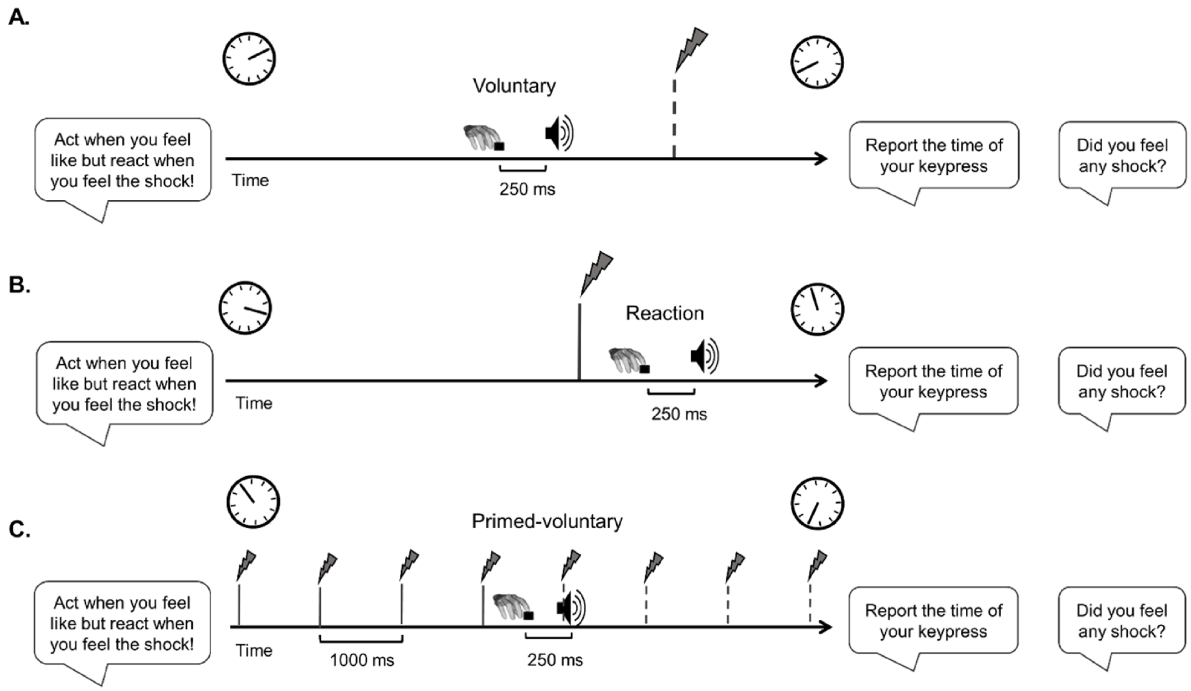


Figure 3.2. Timeline of an experimental trial. Participants were instructed to look at a rotating clock and to press a key at a time of their free choice or to react as soon as possible if they felt a shock. In *voluntary* trials participants pressed the key before occurrence of a

supraliminal shock (large shock sign) (A). In *reactive* trials they pressed the key immediately after feeling the supraliminal shock (the supraliminal shock happened at a random time drawn from an exponential distribution) (B). In *primed-voluntary* trials a subliminal shock (small shock sign) was delivered every 1 s till participants pressed the key (C). Each keypress was followed by a beep 250 ms later. At the end of each trial participants reported the time of their keypress and whether they received a shock. Dashed lines show hypothetical time of a shock.

Trials containing a train of subliminal shocks were also divided into two categories. First, if the participant reported perceiving any shock, the trial was discarded. If the participant did not report perceiving any shock, the trial was categorized as a *'primed-voluntary'* trial.

The baseline condition was tested in two separate blocks of 15 trials each, at the beginning and end of the experiment. The agency condition was tested in four blocks of 40 trials each between the two baseline blocks.

3.2.6. Data analysis.

In signal detection task the proportion of hits, correct rejections, misses and false alarms were calculated separately for supra- and subliminal shocks. These measures where then used to compute the sensitivity index (*d'*).

In the intentional binding task, judgment error was defined as the difference between the judged clock hand position and the actual time of the keypress on each trial. A positive judgement error indicated a perceptual delay; a negative judgement error an anticipation. The mean and standard deviation of the judgement errors across trials were then measured for each trial type. Action binding was defined as the shift of reported time of action towards its outcome, and was calculated by subtracting each participant's mean judgement error in the baseline condition from that in the agency condition. Thus, perceptual association of an action with a subsequent tone would produce a positive value for action binding. We then used repeated-measures ANOVA and paired-samples t-test to compare action binding in *voluntary* trials with action binding in *primed-voluntary* trials. Multilevel models were used

when comparing conditions with unequal sample size, using the *lme* function in R (R Core Team, Vienna, Austria). We did not analyse time judgements from *reactive* trials with supraliminal shocks. The main purpose of having supraliminal shocks was to establish a stimulus-response association between the shock and the action. We reasoned that this makes the shock meaningful for action, and therefore more likely to prime action processing. Finally, a Crawford test (Crawford et al., 2010) was used to compare TP's action binding scores from the healthy and affected hand with the action binding data in healthy participants.

We additionally checked whether subliminal shocks could influence behaviour, as well as sense of agency. The latency of each keypress from the immediately preceding subliminal shock was measured. These latencies were averaged across all *primed-voluntary* trials within each participant. We tested the null hypothesis that the action latencies in *primed-voluntary* trials are from a population with uniform distribution by using a separate Anderson-Darling test for each participant.

3.3. Results

3.3.1. Experience of agency in healthy participants.

Of the 47 recruited participants, 27 met the *d'* criteria of the signal detection task and went on to do the intentional binding task. Four participants could not finish the intentional binding task. Therefore, the final sample included 23 participants (16 females, mean age = 22.7, SD = 3.9). The average detection threshold was 0.5 mA (SD = 0.18 mA). The average supra- and subliminal shock levels were 0.64 mA (SD = 0.23 mA) and 0.49 mA (SD = 0.18 mA), respectively. The average *d'* for subliminal shocks was 1.03 (SD = 0.24). All participants had a d' >= 3 for supraliminal shocks (supplementary table 3.1).

To investigate whether influencing precursor signals to a voluntary action with a subliminal probe could be reflected in one's experience of agency, we compared action binding in

primed-voluntary trials and *voluntary* trials. The perceived time of action moved towards its outcome in both *primed-voluntary* (M = 32 ms, SEM = 7.60 ms, one-sample, t(22) = 4.18, p < 0.01, 95% CI [16 47]) and *voluntary* trials (M = 18 ms, SEM = 8.17 ms, one-sample, t(22) = 2.24, p = 0.03, 95% CI [1 35]) (supplementary table 3.2). However, this action binding was significantly stronger on trials with a subliminal shock train than on trials without shocks (t(22) = 2.61, p = 0.016, d_z = 0.54, 95% CI [3 24]) (Fig. 3.3.A). This suggests that experience of agency towards an action and its effect is associated with precursor brain signals for that action. Finally, to make sure that unbalanced number of trials is not confounding the results, the prevalence of *voluntary* and *primed-voluntary* trials was compared. We found no significant difference (t(22) = 1.54, p = 0.14, 95% CI [-2 15]).

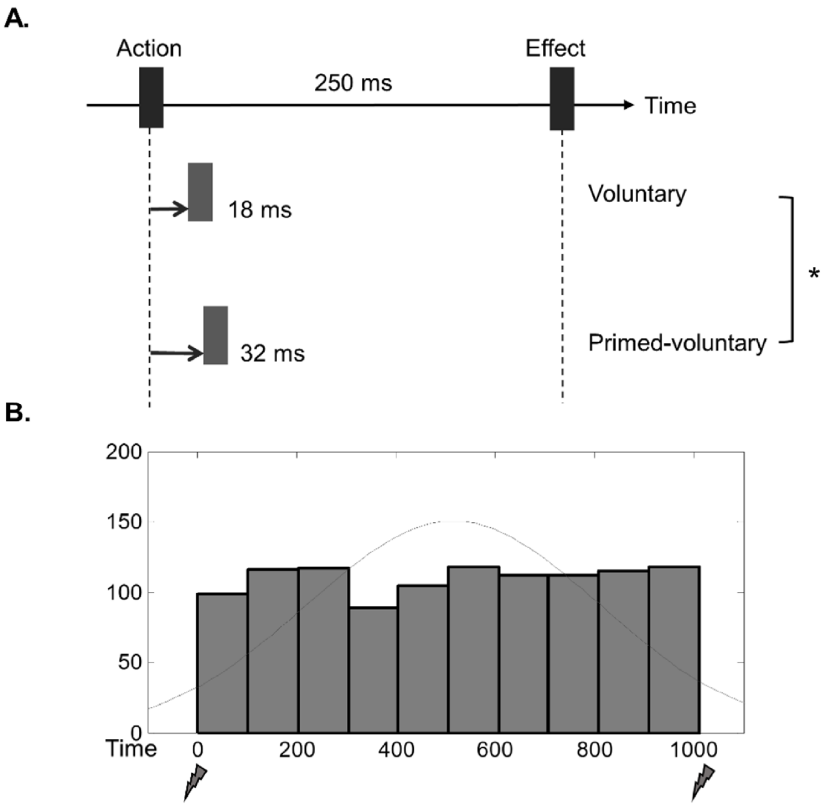


Figure 3.3. Data of healthy participants. A. Action binding in *voluntary* and *primed-voluntary* trials. * *p*<0.05. B. Time histogram of latency of actions from their preceding subliminal shock (time 0), averaged across all *primed-voluntary* trials and all participants. Binding effects are drawn to scale and all values are in ms.

If subliminal shocks influence brain processes during action preparation, we might expect to find the effects not only on *experience* of agency but on some other behavioural measure such as action initiation. We therefore tested the hypothesis that the subliminal shocks influenced the latency of keypresses, by using the Anderson-Darling test to compare keypress latency on *primed-voluntary* trials to a uniform random distribution. The action latency distribution was significantly non-uniform in seven participants (supplementary table 3.3). By the binomial distribution, the probability of this many tests being significant by chance alone is B (0.05, 7, 23) = 0.00009401 (Fig. 3.3.B). This suggests that subliminal shock has some influence on behaviour. However, the presence and pattern of this effect differed across participants. While in some participants subliminal shocks facilitated action initiation, in others it delayed the time of the action (supplementary figure 3.1).

3.3.2. Experience of agency in an individual with anarchic hand syndrome.

TP was tested in two separate sessions, one session for the right (healthy) hand and the other for the left (affected) hand. Detection threshold in the first and second sessions was 0.65 mA and 0.84 mA, respectively. *d'* for subliminal shocks in the first and second sessions was 0.80 and 0.68, respectively.

When testing the healthy hand, perceptual time of action moved towards its outcome in both *voluntary* trials (M = 58 ms, SEM = 23 ms, one-sample, t(28) = 2.54, p = 0.017, 95% CI [11 105]), and *primed-voluntary* trials (M = 100 ms, SEM = 21 ms, one-sample, t(36) = 4.82, p < 0.01, 95% CI [58 142]) (Fig. 3.4.A). In the second session, when testing the affected hand, actions did not bind to their outcomes in *voluntary* trials (M = 35 ms, SEM = 20 ms, one-sample, t(31) = 1.76, p = 0.088, 95% CI [-6 77]), or *primed-voluntary* trials (M = -45 ms, SEM = 40 ms, one-sample, t(23) = -1.13, p = 0.27, 95% CI [-127 37]) (Fig. 3.4.A). Given the unequal number of trials in each condition, factorial repeated-measure ANOVA was performed in a multilevel model with the within subject factors of hand (healthy vs. affected) and trial type (voluntary vs *primed-voluntary*). We found a significant main effect of hand

($X^2(6) = 9.66$, $p < 0.01$), but no significant main effect of trial type ($X^2(7) = 0.31$, $p = 0.58$). Importantly, the interaction between hand and trial type was significant ($X^2(8) = 5.80$, $p = 0.016$). Post-hoc analysis with Wilcoxon signed-rank test showed that the difference in action binding between the two hands was due to the *primed-voluntary* trial types ($p = 0.038$), not the *voluntary* trials ($p = 0.84$) (Fig. 3.4.A & supplementary table 3.4).

The time histogram of latency of keypresses from their preceding subliminal shock in *primed-voluntary* trials is shown for the healthy (Fig. 3.4.B) and the affected (Fig. 3.4.C) hands of TP. Based on the Anderson-Darling test, the distribution of action latencies was not significantly different from a uniform distribution, in the healthy or the affected hand ($p > 0.1$). This finding, however, should be considered in the face of low number of trials from a single case.

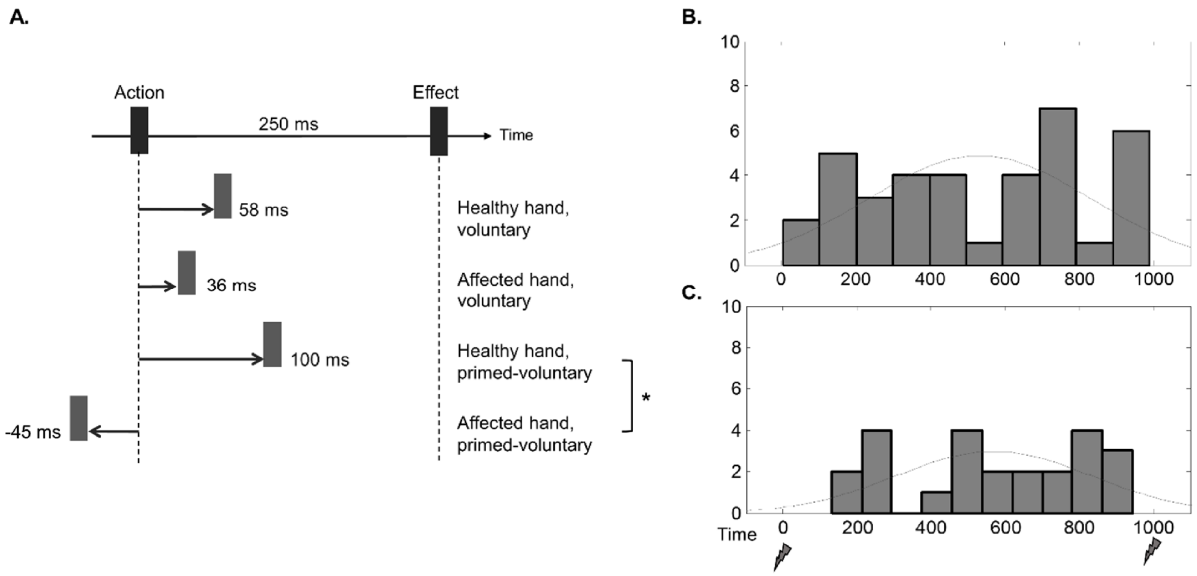


Figure 3.4. Data of TP. A. Action binding in *voluntary* and *primed-voluntary* trials for the healthy and the affected hand. * $p<0.05$. B&C. Time histogram of latency of actions from their preceding subliminal shock, averaged across all *primed-voluntary* trials, displayed separately for the healthy (B) and the affected hand (C). Binding effects are drawn to scale and all values are in ms.

3.3.3. Experience of agency in TP vs. healthy participants.

Finally, we tested whether subliminal shock effects on action binding were significantly different in TP and healthy participants, using Crawford test. This method tests whether a single patient's score differs significantly from that in a control group, and also provides a point estimate of the separation between the patient's score and the control group (Crawford et al., 2010). The effect of the subliminal shocks was measured by subtracting each participant's action binding in *primed-voluntary* trials from *voluntary* trials. The effect of subliminal shock on experience of agency, as measured by action binding, did not differ significantly between healthy participants and the healthy hand of TP ($t = 1.14$, $p = 0.27$, $Z_{cc} = 1.16$). However, while subliminal shocks enhanced action binding in healthy participants (subliminal shock effect = 13 ms), they reduced it in the affected hand of TP (subliminal shock effect = -80 ms), hence showing an opposite effect ($t = -3.64$, $p < 0.01$, $Z_{cc} = -3.72$).

3.4. Discussion

Healthy subjects and an individual with anarchic hand syndrome were exposed to subliminal electrocutaneous stimulus during the precursor period before performing internally-generated actions that produced an external outcome. We used an established implicit measure based on time perception to measure sense of agency. The perceived time of an action has been found to shift towards its outcome for voluntary actions but not for involuntary movements (Haggard et al., 2002b). Using this 'intentional binding' index, we developed a new paradigm to investigate how sense of agency might be influenced by external subliminal stimuli. On one view, such stimuli might influence internal precursors of voluntary action, which in turn influence sense of agency. We investigated whether these subliminal shocks might influence sense of agency by boosting a putative '*internal volitional signal*'. On another view, sense of agency is based only on reconstructive inferences about perceptual events associated with action and outcome. Since the shocks were not

perceived, this model cannot readily explain any effect of shock on sense of agency measures.

3.4.1. Subliminal primes boost sense of agency in healthy participants.

The perceived time of endogenous actions moved towards their outcomes in both *voluntary* and *primed-voluntary* trials. Crucially, action binding was significantly stronger in *primed-voluntary* trials where actions were preceded by a subliminal shock, compared to when they were not. The direction of the effect, shifting action perception towards the subsequent outcome, rules out explanations based on P-centre phenomena (Morton et al., 1976), or anchoring effects of the preceding shocks on time perception. Further, as participants could not feel the subliminal shocks, this difference is unlikely to reflect a conscious decision to control actions in a different way. Most importantly, the difference in action binding between trial types could not easily be explained by a purely post-hoc inference account of sense of agency, since the events perceived are identical in both conditions.

Previous studies showed that explicit agency judgements could be modulated by using visual subliminal priming (Chambon and Haggard, 2012; Haggard and Chambon, 2012; Chambon et al., 2014b). Participants reported stronger experience of agency over action effects when the subliminal prime was compatible, compared to incompatible, with the selected action (Wenke et al., 2010). In those studies, as in our experiment, the prime influenced a stage of action preparation that necessarily precedes both action and its effect. This suggests that sense of agency cannot be purely retrospective. Rather sense of agency must depend, at least in part, on signals arising during action preparation. Of course, this does not rule out a further contribution from retrospective inference.

Additionally, given that subliminal shocks *increased* our measure of sense of agency, external stimulation *facilitated* putative precursor signals during action preparation. At first sight, this may seem paradoxical, given the traditional dichotomy between brain systems underlying internally-generated and externally-triggered actions (Passingham et al., 2010a).

However, substantial cross-talk between the two systems exists. In one study, the reaction time to an external-trigger stimulus was reduced in the very final phases of preparation of a voluntary action (Obhi et al., 2009); also see: (Hughes et al., 2011). We speculate that during action preparation, the subliminal shock is taken as an additional environmental cue. The subliminal shock may "nudge" the signal that generates voluntary action, facilitating a threshold crossing event (Schurger et al., 2012). In our paradigm, participants also occasionally reacted to supraliminal shocks. It remains unclear whether this prior association between shock and action is essential for the subliminal priming we observed. We hope to investigate this point in future experiments. Interestingly, we also found some statistical evidence for effects of shock on action initiation. However, this effect was not present in all participants, and the pattern of influence differed across participants. While in some participants subliminal shocks transiently facilitated action initiation, in others it delayed the time of the action. We note that inhibitory, as well as excitatory, time-dependent effects of subliminal shocks have been widely reported (Blankenburg et al., 2003). We speculate that subliminal shocks may not only sum with the precursor signals during action preparation but also change the threshold for the initiation of the voluntary action. The precise moment of action initiation thus depends on both signal amplitude and the current threshold.

3.4.2. Subliminal primes reduce sense of agency in an anarchic hand.

Patients with AHS often complain of lack of agency for movements made by their affected hand. This was reflected in action binding data from the left (affected) hand of TP. While she perceived the time of the endogenous actions that were performed by her right (unaffected) hand as shifted towards their outcomes, this perceptual shift did not happen for endogenous actions of her affected hand. This finding based on our implicit measure of sense of agency is also in line with TP's subjective reports of episodes of lack of control of her left hand (see section 2.1).

Interestingly, the significant interaction between hand and trial type showed that subliminal shock enhances sense of agency similar to healthy participants, but only when applied to the

healthy hand. Subliminal shock had no statistical effect when applied to the affected hand. We suggest that, for the affected hand, a mechanism that uses precursor signals of voluntary action to compute sense of agency is now disrupted. The normal function of this mechanism would include integrating signals from the external environment and from internal states to construct a coherent subjective experience of action.

Normal behaviour is an outcome of active interplay between internal states and the external environment. Successful interaction of these two components is crucial for goal-directed behaviour and inhibition of unwanted responses. Patients with focal damage in medial frontal cortex (though without signs of AHS) show disruption to automatic motor inhibition, as evident in a reversal of the normal negative compatibility effect in a masked-prime task (Sumner et al., 2007). Abnormal facilitation by priming, as well as the involuntary object-oriented actions that characterise AHS, could both be viewed as productive symptoms reflecting damage to a brain system that normally inhibits excessive environmental reactivity. Our results suggest a second aspect to AHS. The normal subjective experience of action is altered in AHS, and in particular the capacity to feel a sense of agency for voluntary actions that are appropriately interfaced to subtle cues in the external environment.

The brain lesions of TP mostly involved the right anterior cingulate cortex (ACC) and the posterior part of corpus callosum (CC) (Figure 3.1 & Table 3.1). Lesions in these areas have been previously reported in patients with AHS (Hassan and Josephs, 2016). One fMRI study compared brain activity during alien hand and voluntary movements of a patient with AHS (Assal et al., 2007). While alien hand movements were associated with isolated activity in contralateral motor cortex, voluntary movements of the same hand activated extensive networks including the ACC, suggesting a possible role of ACC in voluntary action control. Moreover, ACC has been shown to be active during self- and external-agency attribution tasks (Nahab et al., 2011; Fukushima et al., 2013). Other case studies have associated lesions in the CC with volitional disorders of AHS (Della Sala et al., 1991; Feinberg et al., 1992). CC connects the frontal and motor areas of the two hemispheres. Specifically, the

body and splenium of CC, which are mainly damaged in TP, connect the premotor areas (Berlucchi, 2012). Damage to this area could thus lead to loss of transcallosal motor inhibition of the contralateral hemisphere (Kim et al., 2014). Interestingly, (Wolpe et al., 2014) found a relation between CC white matter loss and abnormal intentional binding in patients with alien limb due to corticobasal degeneration. This deficit was largely confined to anterior parts of CC.

Patients with AHS commonly report that their hand is not under their control or being controlled by an external agent (e.g., Assal, Schwartz, & Vuilleumier, 2007). Our work suggests that this phenomenology may arise from two distinct sources. The first source, and the only one recognised in the current literature, is the positive symptom of the affected hand's performing undesired movements in response to the external world. We suggest here a second source of AHS phenomenology, namely a reduced sense of agency for one's own voluntary actions. In the normal brain, voluntary actions do not come "from nowhere", but are aligned to subtle action possibilities suggested by the environment, akin to subliminal priming in laboratory experiment. Such priming increases explicit judgements of agency (Wenke et al., 2010), and increased intentional binding in healthy volunteers. However, this mechanism was absent for the affected hand of our AHS patient. To our knowledge, this is the first study to investigate a negative symptom of AHS by measuring the effect of the external world on experiences of voluntary actions.

3.4.3. Sense of agency as a readout of internal volitional signals: a cognitive model.

Based on our findings from healthy participants and TP we propose a cognitive model of the experience of voluntary action (Fig. 3.5). We suggest that one key input to the experience of agency is a readout of an internal volitional signal that precedes endogenous actions. This internal signal, however, could be influenced by externally-triggered signals from the outside environment (affordances): volition is not independent of the current environment and response space (Schüür and Haggard, 2011). In the case of healthy participants and the

unaffected hand of TP, this external signal is integrated into the internal volitional signal to facilitate action preparation. Thus, the weak sensory evidence suggesting action that is provided by a subliminal prime is summed with the intention or predisposition to act provided by the task instruction. This integration is accordingly reflected in a stronger action binding and an altered distribution of acting. Thus, suggestions of the external environment are integrated with intentions, and the sense of agency depends partly on a metacognitive readout from the output of this 'integrator' (Fig. 3.5, node 1) (Fleming and Frith, 2014).

This interface between the will and the external world is damaged in AHS (Fig. 3.5). Classical descriptions of AHS suggest that intentional control no longer inhibits affordance-based responding – resulting in compulsive or utilisation behaviours (Fig. 3.5, node 2) (Riddoch et al., 1998; McBride et al., 2013). The affected hand sometimes reacts to the external world due to loss of the normal inhibitory signal of volition (Sumner et al., 2007). Accordingly, the patient's experience of actions is no longer driven by metacognitive readout of one's own intentions, but is instead driven by experience of actual motor outputs triggered by environmental stimuli. As a result, patients with AHS frequently describe movements of the affected hand as involuntary, even when they are well-formed and co-ordinated. For example, patients may report that their affected hand 'has a mind of its own', 'is being 'naughty', 'doing what it wants, not what I want', etc.

This model contains the inhibitory link from the voluntary to the reactive motor system that is classically associated with AHS (Fig. 3.5, node 2). Our results here suggest that the interface also involves a second link, whereby the external environment, even in mild subliminal form, can gently nudge volition (Fig. 3.5, node 1). This nudge can lead to changed behaviour, as in subliminal priming (Eimer and Schlaghecken, 2002), but also changed experience of volition, as in the altered sense of agency here. Damage to the interface area in AHS also weakens this second facilitatory link between the external environment and volition, preventing the normal subliminal facilitation of sense of agency. Taken overall, a healthy sense of agency requires that the voluntary motor system be responsive to

appropriate external suggestions when these align with one's own wishes, while retaining the ability to suppress externally-driven actions when these are not desired. Our results suggest that the cingulate and the callosum participate in this bidirectional interaction.

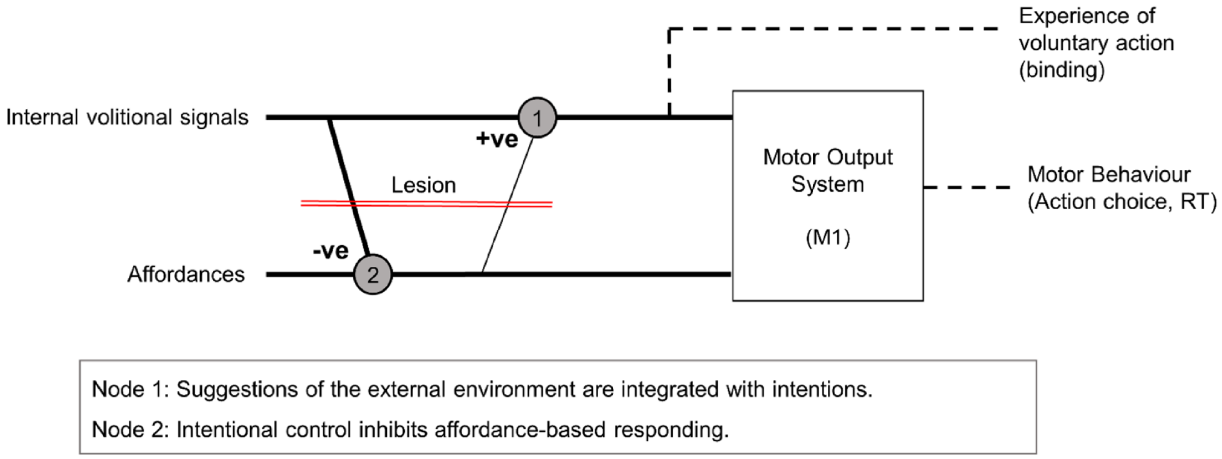


Figure 3.5. A cognitive framework for experience of voluntary action in healthy participants, and anarchic hand syndrome.

3.5. Conclusions.

We developed a novel paradigm to investigate the contribution of precursor signals of endogenous actions to sense of agency. We showed that experience of agency is a metacognitive readout of an interaction between internal volitional signals and the outside world, and not merely a post-hoc confabulation. Interestingly, this interaction was impaired in a patient with anarchic hand syndrome. These findings may help us better understand the mechanisms of volition and sense of agency and to better characterise the neurological disorders of volition.

Chapter 4.

Transcranial direct current stimulation over left angular gyrus decreases the perceptual association between actions and outcomes

We performed three experiments using transcranial direct current stimulation (tDCS) to modulate brain circuits involved in control of action, while measuring stimulation-induced changes in one implicit measure of sense of agency, namely the perceived temporal relationship between a voluntary action and tone triggered by the action. Participants perceived such tones as shifted towards the action that caused them, relative to baseline conditions with tones but no actions. Actions that caused tones were perceived as shifted towards the tone, relative to baseline actions without tones. This 'intentional binding' was diminished by anodal stimulation of the left parietal cortex (targeting the angular gyrus (AG)), and, to a lesser extent, by stimulation targeting the left dorsolateral prefrontal cortex (DLPFC), (Experiment 1). Cathodal AG stimulation had no effect (Experiment 2). Experiment 3 replicated the effect of left anodal AG stimulation for actions made with either the left or the right hand, and showed no effect of right anodal AG stimulation. The angular gyrus has been identified as a key area for explicit agency judgements in previous neuroimaging and lesion studies. Our study provides new causal evidence that the left angular gyrus plays a key role in the perceptual experience of agency.

4.1. Introduction

Despite extensive theoretical work on agency, its neural correlates are not fully understood. Neuroimaging studies found activation of AG (Farrer and Frith, 2002; Farrer et al., 2003a, 2008) and DLPFC (Fink et al., 1999) associated with agency tasks, but the activation of these areas was always greater in the conflicting, non-agency condition than in the agency condition. In a recent meta-analysis of sense of agency, the single most consistent result was activation of a broadly-defined temporoparietal junction area conditions associated with reduced or absent sense of agency (Sperduti et al., 2011). This broad 'non-agency' area includes AG. Computational models of predictive motor control offer an important theoretical framework for understanding agency. An internal forward model uses efference copies of the motor command to predict outcomes (Wolpert and Ghahramani, 2000). According to these models, sense of agency arises when there is a match between the predicted and actual sensory outcome of the generated action. Conversely, if current sensory information does not match the model's prediction, then the corresponding sensory event cannot be self-generated, and no sense of agency is experienced (Frith et al., 2000).

Farrer et al. (2008) used this framework to interpret fMRI activations of AG in particular, suggesting that AG processes discrepancies between intended action and its actual consequences. Her data showed increased activations of AG when a detectable temporal discrepancy was inserted between an action and visual feedback of the outcome, and also when participants explicitly rejected agency over the viewed outcome.

Most of these studies used explicit agency attribution tasks, in which participants *judge* whether they did or did not cause a specific sensory event. (Synofzik et al., 2008) noted that one *feels* a sense of agency when acting, even without making any explicit judgements. One suitable measure of this pre-reflective, sensorimotor feeling of agency is the perceived temporal relationship between a voluntary action and its sensory outcome (Moore and Obhi, 2012). The perceived time of voluntary actions and their sensory consequences are attracted towards each other. This 'intentional binding' is absent, or less prominent, for

involuntary movements, and for associations between external events not involving voluntary actions (Cravo et al., 2009).

The neural bases of such *feelings* of agency are poorly understood. One neuroimaging study found a neural correlate of intentional binding in the medial frontal cortex (Kühn et al., 2013). A 'virtual lesion' study showed that theta-burst stimulation over a slightly more anterior medial frontal location reduced the intentional binding effect (Moore et al., 2010). On the other hand, other lesion (Sirigu et al., 2004) and stimulation (Desmurget et al., 2009) studies suggested an important role of parietal cortex in intentional action and agency, though these studies did not use binding. To our knowledge, no previous causal study has investigated the influence of both frontal and parietal areas on sense of agency using implicit measures. We therefore performed three transcranial direct current stimulation (tDCS) experiments, to modulate excitability of key brain circuits underlying the control of action, while measuring the effects on sense of agency, using intentional binding. Our experiments investigated the respective contributions of parietal and frontal areas to intentional binding as a proxy measure of agency (Experiment 1), their susceptibility to both up- and down-regulation (Experiment 2), and their hemispheric specialisation (Experiment 3).

Based on the existing neuroimaging data investigating explicit agency judgement (Farrer and Frith, 2002; Farrer et al., 2003a, 2008), we predicted that anodal stimulation of putative AG should also influence the sense of agency, as measured by intentional binding. Importantly, such a result would identify a causal role for AG in sense of agency, but would not conclusively identify *how* AG computes agency. We also investigated the role of prefrontal areas in sense of agency. Studies of frontal contributions to sense of agency are more equivocal. Neurostimulation (Moore et al., 2010) and neuroimaging (Kühn et al., 2013) studies of intentional binding found evidence for medial prefrontal involvement, but studies of explicit agency judgements in tasks requiring a choice between alternative actions (Chambon et al., 2013) identified a more lateral prefrontal focus. DLPFC has also been identified as a key area for initiation (Jahanshahi et al., 1995) and monitoring of voluntary

action (Rowe et al., 2010). Given the relative inaccessibility of medial prefrontal cortex to neurostimulation, we focussed here on the lateral prefrontal cortex. The stimulations targeted primarily the left hemisphere, and participants made actions with their right hand (experiments 1,2), or with either hand (experiment 3).

4.2. Materials and Methods

4.2.1. Participants

In total 55 healthy volunteers, 18-35 years of age (25 females) were recruited from the Institute of Cognitive Neuroscience subject data pool for three separate experiments. All participants were right handed, had normal or corrected to normal vision, had no history or family history of seizure, epilepsy or any neurologic or psychiatric disorder and did not have any metallic or electronic object in the head. Participants affirmed that they had not participated in any other brain stimulation experiment in the last 48 hours, nor had consumed alcohol in the last 24 hours. The sample for Experiment 1 consisted of 18 participants (8 females), Experiment 2 consisted of 19 participants (10 females) and Experiment 3 consisted of 18 participants (7 females). One participant failed to finish Experiment 2 due to lack of concentration, and was therefore excluded. Experimental design and procedure were approved by the UCL research ethics committee, and followed the principles of the Declaration of Helsinki. Transcranial stimulation followed established safety procedures (Nitsche et al., 2003; Poreisz et al., 2007). Participants were paid a minimum amount for participating in each session of the experiment. Participants were paid a small additional bonus at the end of the last session.

4.2.2. Behavioural task

We used intentional binding paradigm as an implicit measure of agency (Fig. 4.1). The task was based on previous studies (Haggard et al., 2002b), and was programmed in LabVIEW 2012 (Austin, Texas). Participants viewed a clock hand rotating on a computer screen which was located 60cm in front of the participants in a quiet room. The initial clock position was

random. Clock rotation was initiated by participants pressing the return key on a keyboard. Each full rotation lasted 2560ms. Participants were instructed to look at the centre of the clock. They made voluntary actions, when instructed, by pressing the return key with their right index finger (Experiments 1, 2), or by pressing F9 or F4 with their right or left index finger, respectively (Experiment 3). Participants chose for themselves when to make these voluntary actions. After each key press, the clock hand stopped at a random location, participants made a time judgement according to condition (see later). Each experimental session consisted of four types of trials, presented in separate blocked and randomised conditions. At the beginning of each block, brief instructions for the relevant condition were displayed on the screen. In the *baseline action* condition, participants had to press the key at a time of their own free choice. The clock hand stopped after 1500-2500ms (at random), and participants then judged the clock hand position at the time of their key press, entering their response on the keyboard. In this condition, the participant's actions produced no sensory outcome. In the *baseline tone* condition, participants were instructed to look at the clock but not to press any key. While the clock was rotating, a pure tone (1000Hz, 100ms duration) was played over a loudspeaker, 1750-4000ms (at random) after the onset of the trial. Participants were then asked to judge the clock hand position at the time of the tone. In the *operant action* condition, participants pressed a key at a time of their own choosing, and each keypress produced a tone after 250ms. Participants judged the clock hand position at the moment of pressing the key. Finally, the *operant tone* condition was similar to the operant action condition, with the difference that participants had to judge the clock hand position at the time of the tone. Each condition was tested in a separate block of 30 trials. The order of the blocks was randomised and there was a 1 minute break between each block.

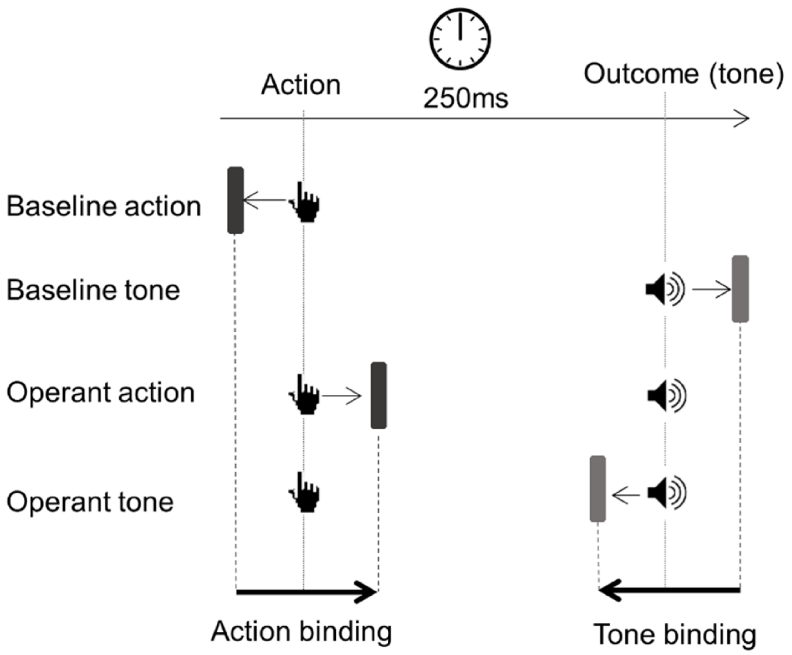


Fig. 4.1. Schematic of intentional binding. Action and tone shifts are measured by subtracting each participant's mean judgement error in baseline conditions from judgement error in operant conditions. These shifts serve as measures of intentional binding. Vertical bars and thin arrows represent mean judgement errors in each condition. Thick arrows represent binding effects. See text for full explanation.

This common basic design was slightly changed according to the demands of each specific experiment. In Experiment 3, text below and above the clock instructed participants to reply with either their left or right index fingers. The order of hands was randomised in each block. The block length was increased by 40 trials to allow sufficient trials for analysis of each hand's data.

Before each experiment, participants were trained and familiarised with the task. They were reminded to look at the centre of the clock, to avoid following the clock hand with their eyes, to be spontaneous in their key presses and to be as precise as possible in their judgements, in particular not confining themselves to those numbers 5,10,15… marked on the clock face. Each experimental session consisted of four blocks and took approximately 20 minutes. The short duration of each individual session was planned to coincide with the known effective period of tDCS.

4.2.3. tDCS

Direct current stimulation was delivered by StarStim noninvasive wireless neurostimulator (Neuroelectrics, Barcelona, Spain). Circular rubber electrodes (25cm^2) were covered in saline-soaked sponges, installed in a 27 channel neoprene cap, and connected to a wireless current generator. tDCS was then controlled by Neuroelectrics Instrument Controller (NIC v1.2) through a separate computer. Current strength was set at 1mA in all experiments, generating a current density of 0.04mA/cm^2 at the scalp surface. For each experiment, all participants underwent three separate sessions of tDCS, two effective stimulations and one sham session. The order of the sessions was randomised and counterbalanced across participants. There was a minimum of 48 hours (and a maximum of 1 week) between each stimulation session to minimise any potential carry over effects of tDCS (Nitsche et al., 2008b). The duration of stimulation in each session was set at 25 minutes, including 30s to ramp-up and down the stimulating current. For the sham condition, electrical current was only applied during the first and last 30 seconds of the stimulation, so as to induce the same cutaneous sensation as real stimulation, and thus blind the participants as to stimulation condition. During the first 5min of each stimulation, participants were asked to relax on their seats and close their eyes. This delay was designed to allow potential neuro-modulatory effects to build up (Zwissler et al., 2014). Next, participants began the behavioural task while stimulation continued (Fig. 4.2). All participants finished the behavioural task approximately after 20min, the same time as the end of stimulation. In case participants finished the task prior to the end of stimulation they were asked to remain seated until the end of the stimulation. In case the task outlasted the stimulation, they continued to perform the task without further stimulation. The task period never exceeded the stimulation period by more than 2 minutes.

Fig. 4.2 shows tDCS montages of the three experiments. In Experiment 1, the anodal electrode was placed on the left DLPFC (F3 according to the 10/20 international EEG electrode placement) or putative left AG (position P3) (Okamoto et al., 2004; Spitoni et al.,

2013) in separate sessions. During the sham session, the position of the stimulating electrode was counterbalanced between F3 and P3. In all three sessions, the return electrode (cathodal) was placed on the right supraorbital area. For Experiment 2, anode and cathode were placed on the putative left AG in separate sessions while the return electrode was placed on the right supraorbital area. This arrangement was retained during the sham session. Experiment 3 used a biparietal montage (Cohen Kadosh et al., 2010; Hecht et al., 2010). For anodal stimulation of the putative left AG, the anode was placed over P3 and cathode was placed over P4. This arrangement was reversed for anodal stimulation of the putative right AG. For sham stimulation, the anode was pseudorandomly placed either at P3 or P4. After each session participants were asked as part of debriefing if they had experienced any notable effects of stimulation. No effects were reported other than mild tingling sensations localised to the electrodes.

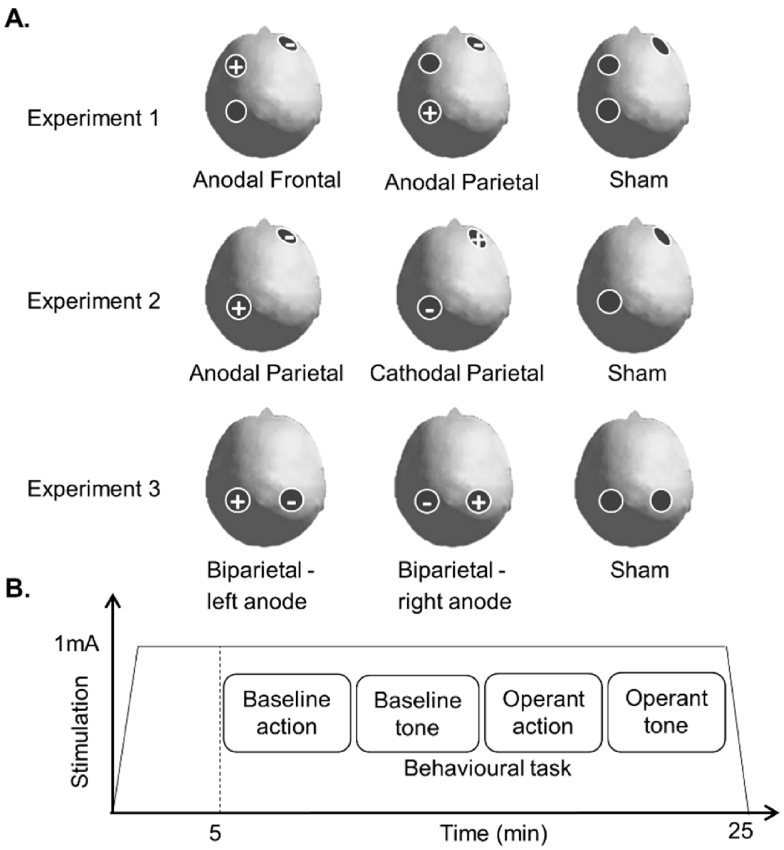


Fig. 4.2.A. tDCS montage and study design (anode +, cathode -). See text for explanation. In Experiment 1, to control for the cutaneous sensation of all three locations, all sponges were kept in place across the three sessions. However, only two of them were actually functioning in each session. B. Stimulation protocol. The order of conditions was randomised within each session.

4.2.4. Data analysis

The difference between the judged clock hand position and the actual onset of the corresponding event was calculated, giving a judgement error for each trial. A perceptual delay was represented by a positive judgement error, and an anticipation by a negative judgement error. The mean and standard deviation of the judgement errors across trials were then measured for each condition. Action binding was defined as the shift of action toward its outcome, and was calculated by subtracting each participant's mean judgement error in the *baseline action* from that in the *operant action* condition. Likewise, tone binding was defined as a shift in the perceived time of a tone towards the action that caused it. Tone binding was calculated by subtracting each participant's mean judgement error in *baseline tone* condition from that in the *operant tone* condition. Thus, perceptual association of an action with a subsequent tone produced a positive value for action binding, and a negative value for tone binding. We analysed action and tone binding separately, since there is evidence from both cognitive studies (Wolpe et al., 2013), and previous neurostimulation studies (Moore et al., 2010), that they are driven by distinct mechanisms.

Some participants were excluded because of highly variable time judgement. A standard deviation of judgement error across trials of over 250ms in any condition was used as a marker of poor time perception. As in previous intentional binding experiments (Haggard et al., 2002a), these participants were excluded. On this basis, two participants were excluded from Experiment 1, two from Experiment 2, and four from Experiment 3. Importantly, these exclusion criteria are orthogonal to the *mean* judgement errors used for statistical inference.

In Experiments 1 and 2, inferential statistics were based on one-way repeated measures ANOVA, with paired-sample t-tests for follow-up testing. Because our ANOVA had only 3 levels, Bonferroni correction was not required for follow-up testing after a significant result (Meier, 2006; Cardinal and Aitken, 2013). Additionally, in Experiment 1, we used linear discriminant analysis to determine which percepts were most strongly affected by stimulation condition. Experiment 3 used repeated measures ANOVA, with the additional factor of acting hand (right hand responses vs. left hand responses). A final pooled analysis was performed to compare effects of stimulation common to all conditions.

4.3. Results

4.3.1. Experiment 1: frontal vs parietal anodal stimulation

This experiment compared the effects of frontal (targeting left DLPFC) and parietal (targeting left AG) cortex stimulation on intentional binding for actions and tones. One-way repeated measures-ANOVA with the factor of stimulation type (anodal frontal vs. anodal parietal vs. sham) showed that action binding was not significantly affected by the type of stimulation ($F(2, 30)=1.90$, $p=0.17$, $\eta^2=0.11$). However, an identical ANOVA on tone binding showed significant differences ($F(2, 30)=4.30$, $p=0.02$, $\eta^2=0.22$). Follow-up testing showed that tone binding was significantly reduced by anodal stimulation of the putative left AG compared to sham ($t(15)=2.67$, $p=0.02$, $d=0.43$). Anodal stimulation of the left DLPFC showed a clear trend to reduce tone binding, which approached the border of conventional significance ($t(15)=2.07$, $p=0.06$, $d=0.42$). There was no significant difference between the frontal and parietal stimulation ($t(15)=-0.10$, $p=0.92$, $d=0.02$) (Fig. 4.3).

We additionally performed the same analysis using median rather than mean judgement error in each condition, since median measures are more robust than means to the influence of outliers. The patterns of statistical significance were unchanged.

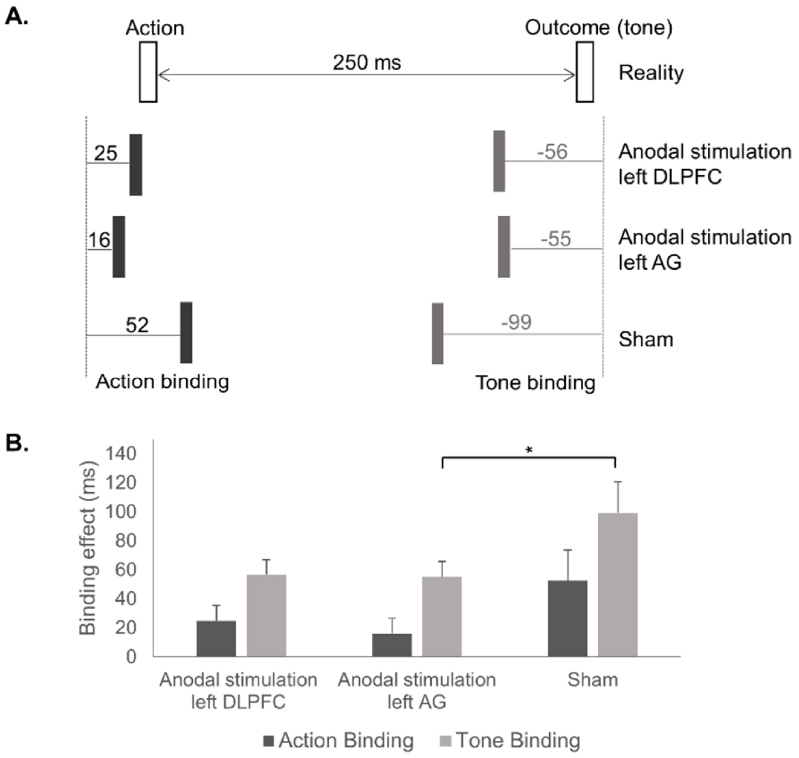


Fig. 4.3. Intentional binding in Experiment 1. A) The dashed line indicates the perceived time of either action or tone in the corresponding baseline condition. A separate baseline condition was used for each session, and differences in baseline values across sessions have been removed for display purposes. Binding effects are drawn to scale, and values are in milliseconds. B) Mean binding effects in ms. The sign of tone binding effects has been inverted to allow for comparison with action binding. Error bars show standard error of the mean. * $p<0.05$.

Finally, to confirm that anodal stimulation of putative AG affected primarily the operant tone condition and not the baseline tone, the effect of stimulation type on participants' judgement error in *baseline* tone conditions was assessed using repeated-measure one-way ANOVA. There was no significant main effect of stimulation type on baseline tone condition ($F(2, 30)=0.58$, $p=0.56$, $\eta^2=0.04$). This suggests that stimulation influenced a neurocognitive process that is present primarily in the operant condition.

We additionally applied multivariate linear discriminant analysis (Krzanowski, 2000) to identify the linear combination of action binding and tone binding variables that optimally discriminates the different stimulation conditions. Linear discriminant analysis significantly differentiated the three stimulation conditions (Wilks' Lambda=0.59, approx. F(4,58)=4.36, $p<0.01$). Inspection of canonical coefficients showed that this difference was primarily due to tone binding (standardized canonical coefficient 1.86) rather than action binding (-0.93) (The scores of the individual participants on the first discriminant variate are shown in Fig. 4.4). Post-hoc comparisons between conditions showed a highly significant difference between parietal and sham stimulation ($p<0.01$; standardised coefficients -1.03 for action binding, 2.19 for tone binding), and also a significant difference between frontal and sham stimulation ($p=0.04$; standardised coefficients -0.82 for action binding, 1.55 for tone binding). Interestingly, the frontal effect thus involved a slightly larger action binding coefficient, considered relative to the tone binding coefficient, than did the parietal effect, though no inferential statistics can be applied to this ratio. Frontal and parietal stimulation did not differ significantly ($p=0.74$).

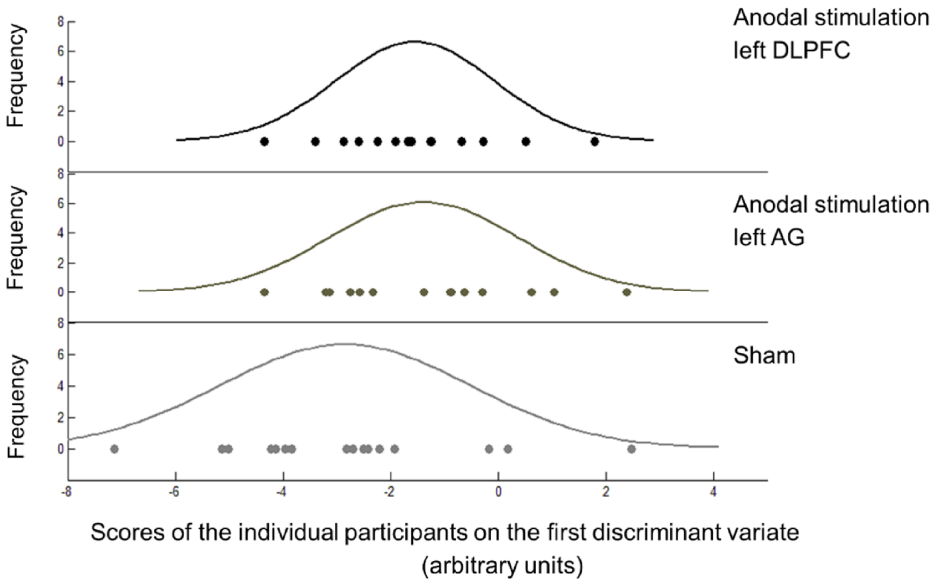


Fig. 4.4. Scores of the individual participants on the first canonical variate from a discriminant analysis of Experiment 1. Each dot indicates one participant, and a fitted Gaussian frequency distribution is also shown. The raw coefficients of the canonical variate were 0.015 for action binding, and 0.02 for tone binding. Both parietal and frontal stimulation conditions were significantly different from sham.

4.3.2. Experiment 2: anodal vs cathodal parietal stimulation

If anodal stimulation boosts activity in the left AG then cathodal stimulation of the same area should lead to its suppression. Thus, if anodal stimulation decreases intentional binding, cathodal stimulation should increase it. To test this hypothesis, putative left AG was exposed to anodal, cathodal or sham stimulation in different sessions, and effects on intentional binding were evaluated.

One-way repeated measures-ANOVA with the factor of stimulation type (anodal parietal vs. cathodal parietal vs. sham) was used for analysis. Action binding was not significantly affected by the type of stimulation ($F(2, 30)=0.50$, $p=0.61$, $\eta^2=0.03$). Tone binding was also unaffected by type of stimulation ($F(2, 30)=0.45$, $p=0.64$, $\eta^2=0.03$), contrary to our predictions from Experiment 1. Nevertheless, the numerical effect of anodal stimulation of the putative left AG was in the same direction as Experiment 1, namely a decreased tone binding compared to sham and cathodal stimulation (Fig. 4.5).

We additionally performed the same analysis using median rather than mean judgement error in each condition. The patterns of statistical significance were unchanged.

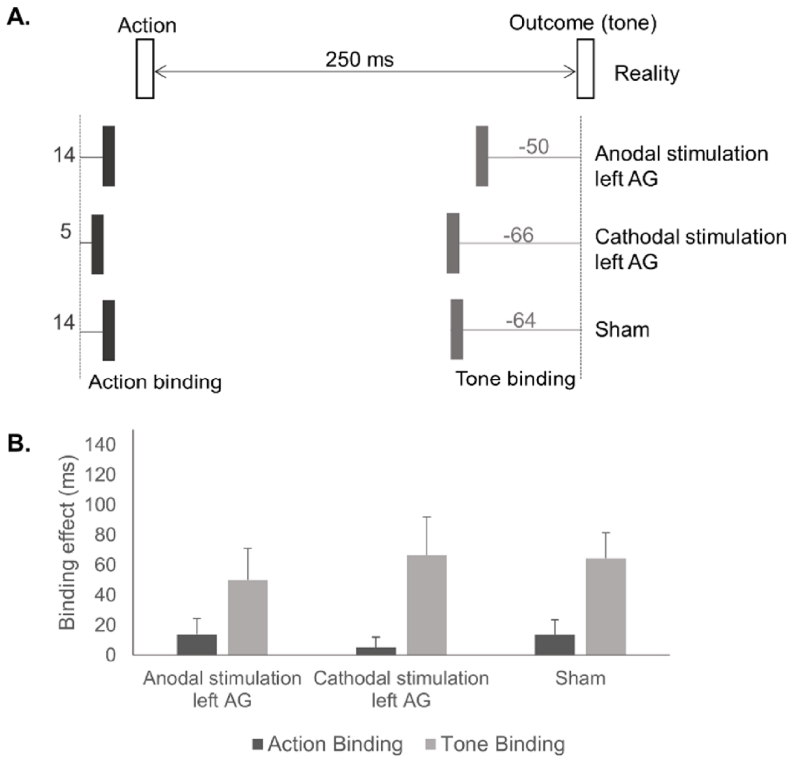


Fig. 4.5. Intentional binding in Experiment 2. Format as in Fig. 4.3.

4.3.3. Experiment 3: left vs right parietal stimulation

Experiment 3 aimed to replicate the effects of Experiment 1, and additionally investigated the lateralisation of intentional binding using a biparietal montage. The biparietal montage may provide a higher local current density, because of the relatively short path between anode and cathode. In addition, this montage controls for any possible effect of cathodal stimulation of prefrontal areas that may occur with the conventional supraorbital placement of the cathode. Therefore, putative left and right AG were exposed to anodal stimulation in separate sessions, and a third session involved sham stimulation. We also investigated whether the putative AG involvement in intentional binding is hand-specific or hemisphere-specific, by asking participants to make actions with either the left or right hand, choosing randomly on each trial. Analysis of action binding showed no significant main effect of

stimulation type ($F(2, 26)=0.06$, $p=0.94$, $\eta^2=0.01$) or acting hand $F(1, 13)=0.10$, $p=0.76$, $\eta^2=0.01$) and no significant interaction ($F(2, 26)=0.89$, $p=0.42$, $\eta^2=0.06$). Analysis of tone binding showed a highly significant main effect of stimulation ($F(2, 26)=5.93$, $p<0.01$, $\eta^2=0.31$). Follow-up testing showed that anodal stimulation of the putative left AG significantly decreased tone binding relative to both sham stimulation ($t(13)=2.55$, $p=0.02$, $d=0.40$), and relative to anodal stimulation of the putative right AG ($t(13)=2.90$, $p=0.01$, $d=0.56$). No significant difference was observed between anodal stimulation of the putative right AG and sham ($t(13)=-0.75$, $p=0.47$, $d=0.10$) (Fig. 4.6). Acting hand had no significant main effect on tone binding ($F(1, 13)=0.01$, $p=0.94$, $\eta^2<0.01$) and no interaction was observed between the stimulation and acting hand ($F(2, 26)=0.15$, $p=0.86$, $\eta^2=0.01$).

We additionally performed the same analysis using median rather than mean judgement error in each condition. The patterns of statistical significance were unchanged.

To check whether the decrease in tone binding was primarily due to shifts in the operant tone, or in the baseline tone condition, participants' judgement errors in the baseline tone condition were compared across the stimulation groups. Analysis showed no significant main effect of stimulation type on baseline tone condition ($F(2, 26)=0.50$, $p=0.60$, $\eta^2=0.04$).

4.3.4. Pooled data: anodal parietal vs sham stimulation

Anodal stimulation of the putative left AG was common to all three experiments reported here, as was a sham stimulation condition, although the experiments differed in other respects. Therefore, we pooled the data in these specific conditions across the 46 participants (21 female) from the three experiments, in a single analysis. We found that anodal stimulation of putative left AG significantly reduced the perceptual shift of tone toward action compared to sham ($t(45)= 3.28$, $p<0.01$, $d=0.35$), but had no effect on action binding ($t(45)=-1.37$, $p=0.18$, $d=0.22$).

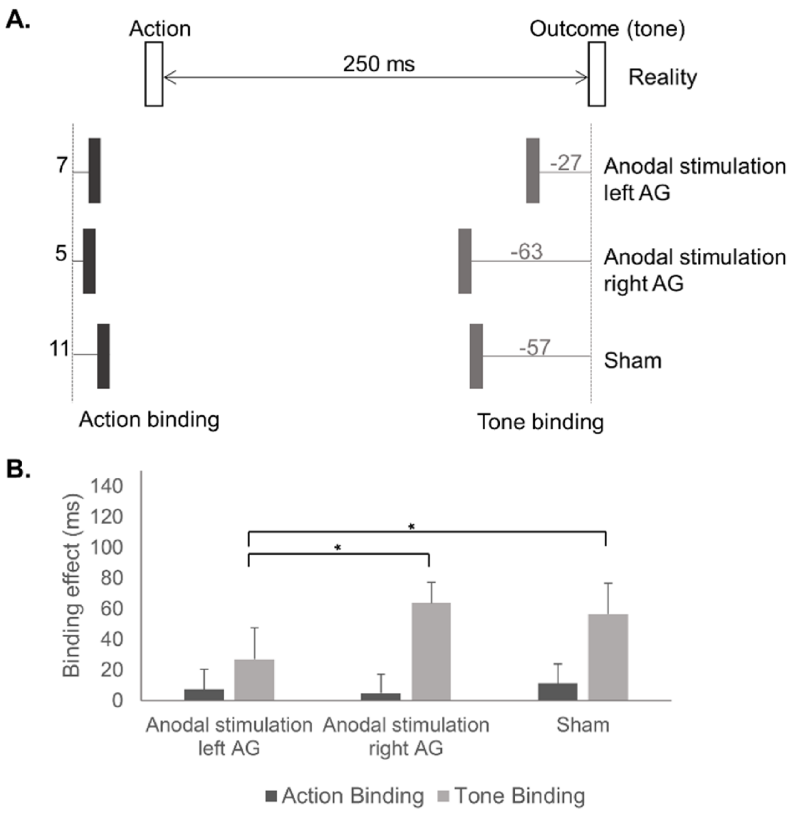


Fig. 4.6. Intentional binding in Experiment 3. Format as in Fig. 4.3.

The anodal montage in the third session used a different return (cathode) location compared to the first and second experiments. Our decision to pool data of the three studies was based on the common placement of the anode across these three experiments, and the general observation that cathodal effects on cognitive function are rare (Jacobson et al., 2012b). In that case, differences in cathode location may be relatively unimportant, and need not prevent pooling across studies. However, because we cannot entirely exclude some contribution of cathodal location to our main results, we ran a further pooled analysis using the left anodal stimulation conditions of experiments 1 and 2 only, which share a supraorbital cathode location, but excluding experiment 3. This analysis again found that anodal stimulation of putative left AG significantly reduced the perceptual shift of tone toward action compared to sham ($t(31)= 2.45$, $p=0.02$, $d=0.35$), but had no effect on action binding ($t(31)=-1.48$, $p=0.15$, $d=0.32$).

To investigate the generality of the anodal AG effect across experiments, we performed a mixed ANOVA with a between-subject factor of experiments (1, 2 or 3), and a repeated-measures factor of stimulation type (anodal left AG vs. sham). The main effect of stimulation type ($F(1, 43)=10.7$, $p<0.01$, $\eta^2=0.20$) recapitulating the pooled t-test reported above. There was no significant main effect of experiment ($F(2, 43)=0.80$, $p=0.45$, $\eta^2=0.04$). Importantly, there was no hint of interaction between experiment and stimulation ($F(2,43)=0.96$, $p=0.39$, $\eta^2=0.04$).

Our main inferences above are based on comparing experimental stimulation with sham. We therefore additionally investigated whether sham stimulation had different effects in the three experiments. We found no significant difference among the sham conditions for the 3 experiments for action binding ($F(2,43)=2.20$, $p=0.12$), or tone binding ($F(2,43)=1.10$, $p=0.33$).

4.4. Discussion

4.4.1. Stimulation-induced modulation of the left parietal cortex

We performed a series of three tDCS experiments to investigate the neural circuits responsible for the sense of agency, as measured by the perceptual association between the time of a voluntary action and the time of a resulting auditory tone. We found a significant decrease in the binding of outcomes towards actions after anodal stimulation of the putative left AG (Experiment 1). Anodal stimulation of the left DLPFC also decreased action and tone binding compared to sham. DLPFC affected tone binding as much as AG stimulation, but its effects were less consistent across participants. Nevertheless, our discriminant analysis showed a significant effect of DLPFC stimulation compared to sham, when action and tone binding were considered together.

Our tDCS stimulation of putative AG could have widespread effects across the inferior parietal cortex (IPC), since tDCS has quite low spatial specificity. For example, anterior parts

of the IPC may also be affected. IPC is routinely activated in neuroimaging studies when participants judge whether their own action, or some other cause, is responsible for a specific sensory event. In a study by Farrer and Frith (Farrer and Frith, 2002), the IPC area was more active when participants attributed a visual event to another person, rather than to themselves. Similarly, a PET study (Farrer et al., 2003a) observed that neural activity in IPC increased with the level of discrepancy between the executed and the observed action on the screen. In an fMRI study (Farrer et al., 2008), the subjective feeling of loss of control correlated with BOLD response in the AG, as did the awareness of temporal discrepancy between action and feedback. The authors of those studies suggested that AG houses the comparison between the efference copy of the intended action and the actual sensory outcome. Any mismatch between these signals will then give rise to the explicit awareness of non-agency, or an external source of action.

Our overall results are consistent with this view. We found that anodal stimulation of the putative AG *decreases* intentional binding, our proxy measure for agency. Anodal stimulation is generally thought to *increase* the activity of the cortical region immediately under the electrode. However, AG activation is routinely associated with *lack of agency*, rather than with experience of positive agency (Farrer and Frith, 2002; Farrer et al., 2008; Sperduti et al., 2011). Therefore, excitation of a neural substrate of non-agency might be expected to *decrease* intentional binding. The conventional polarity-specific (anode-boosting, cathode-suppressing) framework of tDCS was developed on the basis of effects in primary motor cortex stimulation (Nitsche and Paulus, 2000). Its applicability to non-primary areas and cognitive processing has recently been questioned (Horvath et al., 2014). Nevertheless, our findings are broadly compatible with the conventional polarity-specific view.

4.4.2. Stimulation-induced modulation of the left frontal cortex

Since DLPFC is normally thought to *facilitate* intentional action, one may question why prefrontal anodal tDCS did not *increase* intentional binding. (Rowe et al., 2010) questioned

whether DLPFC played any important role in initiation of simple voluntary actions, such as those tested here, and suggested a role in monitoring sequential action patterns instead. (Fink et al., 1999) observed activation of DLPFC using PET when an intentional action and its sensory outcome were incompatible. Anodal tDCS over DLPFC might correspond to an increased coding for action-outcome conflict, even though our task did not explicitly manipulate action-outcome compatibility. Our discriminant analysis found some evidence consistent with this interpretation. However, this effect was investigated in only one experiment, and achieved statistical significance in multivariate analysis, but not in univariate analyses of action binding and tone binding separately. Therefore, further research is required before a strong statement about frontal tDCS effects on sense of agency can be made.

4.4.3. Polarity-specific effects of tDCS

Experiment 2 aimed to investigate whether parietal stimulation effects were polarity-specific. On one model, tDCS would simply add neural noise, irrespective of polarity. On another model, anodal stimulation would upregulate putative non-agency coding in AG, while cathodal stimulation should down-regulate it. The result of anodal left AG stimulation in experiment 2 followed the expected trend for tone binding, but did not reach statistical significance. Replication of statistically significant results is an important and controversial issue in modern neuroscience (Cumming, 2005). All effects measured in experiments represent a combination of the underlying 'true' effect, and noise. Importantly, when a nonzero true effect indeed exists, but is modest in size, it is quite likely for the effect to reach statistically significant levels in one study, but not in another. Thus, absence of a significant anodal tDCS effect in experiment 2 does not prove that no true effect exists: we return to this point later.

Experiment 2 found no significant difference between cathodal and sham stimulation, although we had predicted that cathodal stimulation might enhance intentional binding.

Although inhibitory cathodal effects on motor function are well established, a recent review of 34 studies found that cathodal inhibitory effects on cognitive function are rare (Jacobson et al., 2012b). Another possible reason for the absence of any significant cathodal AG effect in Experiment 2 could be the placement of the anode electrode on the supraorbital area. This location is standard for tDCS studies of action (Nitsche et al., 2008b). However, it causes a strong current density close to the frontopolar and prefrontal areas, where the anode is located. These areas may also contribute to intentional action (Brass and Haggard, 2007). Thus, our montage for cathodal stimulation of AG in experiment 2 involved anodal stimulation at a frontopolar site, which may not be strictly neutral for sense of agency. Future studies could address this issue by using extracephalic cathode placement.

4.4.4. Hemispheric specialisation of the sense of agency

Experiment 3 avoided the potential confound of frontopolar stimulation using a biparietal montage. This produces a higher current density in a small region surrounding the electrodes (Nathan et al., 1993; Cohen Kadosh et al., 2010), compared to the conventional supra-orbital location. This might result in a more focal stimulation. More importantly, the biparietal montage excludes the possibility that the significant effects of anodal AG stimulation in experiments 1 and 2 were in fact caused by cathodal frontopolar stimulation. Specifically, if the effects in experiments 1 and 2 were merely due to cathodal frontopolar stimulation, then no effect of stimulation should be found in experiment 3. The biparietal montage also allowed us to investigate lateralisation of agency by varying both tDCS polarity and the hand used for action. Similar approaches have been used previously in other studies (Jacobson et al., 2012a; Bardi et al., 2013).

The results of the third experiment replicated our previous findings. Anodal stimulation of putative left AG significantly decreased tone binding compared to both sham and cathodal stimulation of the same area. The tDCS effect was statistically equivalent whether the action

was made with the left or the right hand. No effects were observed with anodal stimulation of the putative right AG.

Experiment 3 does not support the alternative interpretation of experiments 1 and 2 based on a putative cathodal frontopolar stimulation. In contrast, experiment 3 supports the interpretation of an anodal left AG effect. We cannot conclusively rule out some contribution of frontopolar stimulation to our results, but we can rule in a specific contribution of the left AG.

Experiment 3 adds several important elements to the previous studies. First, it demonstrates an involvement of AG in a task involving randomised, stimulus-driven selection between alternative actions, as opposed to mere repetition of a simple action. Second, it suggests that left, but not right AG is responsible for action-outcome binding for actions made by either hand. We found no interaction between stimulation and hand used for action. Previous neuroimaging studies have reported activation corresponding to non-agency judgements in both left and right AG. Interestingly, right AG activations appeared to dominate (Farrer and Frith, 2002; Farrer et al., 2003a, 2008), in contrast to our finding. However, in a more recent fMRI study, (Lee and Reeve, 2013) reported higher activity in the left AG during non-self-determined behaviour, consistent with our hypothesis in Experiment 1 that anodal AG stimulation activates a neural code for 'non-agency'. Finally, hemispheric specialisation of agency could plausibly depend on the task used, and the type of agency judgement. Previous neuroimaging studies generally used explicit judgements of agency, and often used complex manual actions with visual feedback (Farrer and Frith, 2002; Farrer et al., 2008; Sperduti et al., 2011). We are not aware of any neuroimaging study investigating the hemispheric lateralisation of low-level implicit measures of agency.

4.4.5. Limitations

The results of experiment 3 by themselves could not distinguish between an effect of anodal stimulation of the putative left AG from an effect of cathodal stimulation of the putative right

AG. However, this result does allow us to exclude a model in which tDCS simply acts to increase neural noise, irrespective of polarity. Moreover, our experiment 1 found some evidence of a left-hemisphere anodal tDCS effect, while our experiment 2 found no evidence of any cathodal effect (though in the left hemisphere, rather than the right). Cathodal stimulation effects in cognitive tasks are reported to be weak (Jacobson et al., 2012b). Therefore, we provisionally favour an interpretation of experiment 3 based on a left parietal anodal effect, rather than a right-hemisphere cathodal effect. Further research would be required to draw a definitive conclusion.

Our study is further limited because we did not control for cases of crossmodal binding in the absence of active movement. Therefore, we cannot exclude the possibility that AG stimulation influenced some general feature of time perception, as opposed to temporal processing specific to agency. However, several studies have shown stronger binding between voluntary actions and outcomes than between other, similarly paired, events, including involuntary movements and outcomes (Engbert, Wohlschläger, Thomas, & Haggard, 2007; Haggard et al., 2002) or pairs of sensory stimuli (Haggard, Aschersleben, Gehrke, & Prinz, 2002; Haggard, Martin, Taylor-Clarke, Jeannerod, & Franck, 2003). Moreover, other studies have investigated effects of parietal tDCS on time perception in general, in the absence of action: and these studies found no effect (Woods et al., 2014). Thus, the weight of other studies suggests that the intentional binding phenomenon reflects a distortion of perceptual timing that is, at least partly, specific to voluntary action.

4.4.6. Dissociation between action binding and tone binding

Anodal stimulation over putative left AG was a common condition in all 3 experiments. Accordingly, we could perform a pooled analysis of intentional binding results to compare this to the sham stimulations that were also included in each experiment. This analysis showed a highly significant reduction in tone binding with anodal stimulation of the putative left AG. We found no overall effect on action binding. Dissociations between action binding

and tone binding have been reported previously (Wolpe et al., 2013), so it is possible that left parietal cortex is concerned primarily with tone binding, rather than with action binding. This conclusion would be consistent with previous studies suggesting that the AG processes mismatches in action outcomes (Farrer et al., 2008). On the other hand, recent studies of explicit agency judgement suggest that AG also processes prospective, premotor information arising during action selection (Chambon et al., 2013, 2014a).

Both online prospective and retrospective processes contribute to the intentional binding phenomenon (Moore and Haggard, 2008). The experimental design used here cannot identify the independent contribution of each process. However, binding of action towards outcome may rely more on prospective processes during action selection, while perceptual shift of outcome toward action may depend on retrospective, more inferential processes, triggered by reafferent signals about action outcome (Chambon, Moore, & Haggard, 2014). Future studies may address this issue by designing new paradigms which dissociate prospective and retrospective components of agency and examine the role of dLPFC and AG in each of these components.

4.4.7. Conclusion and clinical implications

Sense of agency is an important and distinctive feature of human voluntary action. We used a causal intervention (tDCS) and an implicit perceptual measure of sense of agency (intentional binding) to examine the role of different brain areas in sense of agency. Anodal stimulation of parietal cortex consistently reduced the binding of tones towards actions. We hypothesised that the angular gyrus might contribute to the sense of agency by monitoring the linkage of actions to outcomes, or, alternatively and equivalently, failures of such linkage. Anodal stimulation of this area may correspond to artificial boosting of a mismatch detection process.

Sense of agency is altered following several classes of psychiatric and neurological disorders. In particular, patients with apraxia following lesions to the left parietal fail to

recognise the source of a viewed manual gesture (Sirigu, Daprati, Pradat-Diehl, Franck, & Jeannerod, 1999). This deficit is formally equivalent to an overestimation of agency in an explicit judgement task, consistent with damage to a neural centre detecting mismatches. The posterior form of 'alien hand syndrome' is also associated with contralateral parietal lesions (e.g., (Kloesel et al., 2010)). Interestingly, these patients show involuntary and spontaneous movements of the contralateral limb, but may correctly perceive that they are not agents over these actions. The capacity for voluntary movement is often preserved. Quantitative assessment of the implicit sense of agency in parietal patients would be of considerable value in understanding the neural basis of sense of agency.

Chapter 5.

Endogenous action selection processes in dorsolateral prefrontal cortex contribute to sense of agency

Given the role of dorsolateral prefrontal cortex (DLPFC) in action selection, we predicted that DLPFC may contribute to sense of agency when participants select between multiple actions. Ee performed a series of experiments, manipulating a range of task parameters related to action selection and action outcomes, while participants were exposed to tDCS stimulation of the left DLPFC. We measured the temporal association between a voluntary action and its outcome using the intentional binding effect, as an implicit measure of sense of agency. Fixed-effect meta-analysis of our primary data showed a trend towards a frontal tDCS, together with considerable heterogeneity between our experiments. Classifying the experiments into subsets of studies, according to whether participants endogenously selected between alternative actions or not, explained 71% of this heterogeneity. Anodal stimulation of DLPFC increased the temporal binding of actions towards tones in the subset of studies involving endogenous action selection, but not in the other studies. DLPFC may contribute to sense of agency when participants selected between multiple actions. This enhanced feeling of control over voluntary actions could be related to the observed therapeutic effects of frontal tDCS in depression.

5.1. Introduction

In previous chapter we investigated the contributions of parietal and frontal areas to sense of agency by combining tDCS and intentional binding. Anodal stimulation of the left AG reduced intentional binding, in three separate experiments. We suggested that AG may generate experience of agency by monitoring the linkage of actions to outcomes. However, the possible role of frontal cortex is less clear. In one experiment, we found that anodal stimulation of the dorsolateral prefrontal cortex (DLPFC) had a much weaker effect on intentional binding.

DLPFC has been identified as key area in selection and monitoring of voluntary action (Jahanshahi et al., 1995; Rowe et al., 2010). Previous studies of intentional binding focussed primarily on medial frontal cortex (MFC) (Moore et al., 2010; Kühn et al., 2013). The division of labour between these two centres in voluntary action remains unclear, but the DLPFC may be related to selection (Rowe et al., 2000), while the MFC related to the 'urge' or motivation to act (Fried et al., 2011).

Here we focussed on tDCS of the DLPFC. Given previous evidence, we predicted that DLPFC may contribute to sense of agency when participants select between multiple actions. We designed several independent experiments to test whether DLPFC tDCS would alter intentional binding in the context of: endogenously selected actions with different outcome identities (Exp.1), endogenously selected actions converging on the same outcome identity (Exp.2), rule-based endogenously selected actions with different outcome identities (Exp. 3), endogenously selected actions with different outcome values (Exp.4), exogenously-instructed actions converging on a single outcome (Exp.5), and endogenously selected actions with uncertain outcomes (Exp. 6). For purposes of meta-analysis, we also included our previous results from chapter 4, involving endogenous initiation of a single action, producing a single outcome – we refer to this here as dataset 7.

We predicted that left DLPFC tDCS would modulate intentional binding relative to a sham control condition. The overall aim of the series of experiments was to identify if and how the DLPFC might contribute to agency, by investigating a specific experimental condition in each experiment.

5.2. Material and methods

5.2.1. Participants

In total 100 healthy volunteers, aged 18-35 years of age, were recruited from the Institute of Cognitive Neuroscience subject data pool for six separate experiments. No other experiment was performed involving these conditions or measures. All participants were right handed, had normal or corrected to normal vision, had no history or family history of seizure, epilepsy or any neurologic or psychiatric disorder and did not have any metallic or electronic object in the head. Participants affirmed that they had not participated in any other brain stimulation experiment in the last 48 h, nor had consumed alcohol in the last 24 h. Experimental design and procedure were approved by the UCL research ethics committee, and followed the principles of the Declaration of Helsinki.

5.2.2. Behavioural task

The intentional binding paradigm was used as an implicit measure of sense of agency across our experiments (Haggard et al., 2002b). Briefly, in each experimental session participants looked at a clock hand rotating on a computer screen (viewing distance: 60cm). The initial clock position was random. Clock rotation was initiated by participants pressing the return key on a keyboard. Each full rotation lasted 2560 ms. Participants were instructed to look at the centre of the clock and to make time judgement according to the condition. Each condition was presented in a separate and randomised block. Brief instructions were displayed on the screen before each condition. In the two action conditions, participants had to press a key on the keyboard at a time of their own free choice. This keypress either

produced a tone after 250ms (*operant action*) or produced no sensory outcome (*baseline action*). The clock hand stopped after 1500–2500 ms (at random), and participants then judged the clock hand position at the time of their key press, entering their response on the keyboard. For the *baseline tone* condition, participants were instructed to look at the clock but not to press any key. While the clock was rotating, a pure tone (based on the experiment) was played over a loudspeaker, 1750–4000 ms (at random) after the onset of the trial. Participants were then asked to judge the clock hand position at the time of the tone. Finally, the *operant tone* condition was similar to the *operant action* condition, with the difference that participants had to judge the clock hand position at the time of the tone and not the keypress. All four conditions were done by all participants, except experiments 1 and 6 which only involved action binding, and so tested only baseline action and operant action conditions.

This common basic design was modified according to the specific requirements of each experiment (Table 5.1).

5.2.3. tDCS procedure

Direct current stimulation was delivered by StarStim noninvasive wireless neurostimulator (Neuroelectrics, Barcelona, Spain). Circular rubber electrodes (25 cm^2) were covered in saline-soaked sponges, installed in a neoprene cap, and connected to a Neuroelectrics Instrument Controller (NIC v1.2). Current strength was set at 1 mA in all experiments, generating a current density of .04 mA/cm^2 at the scalp surface. A common tDCS montage was used for all six experiments. For anodal stimulation of the left DLPFC, anodal electrode was placed on F3 (according to the 10/20 international EEG electrode placement) and cathodal electrode on right supraorbital area. This arrangement was reversed for the cathodal stimulation of left DLPFC. For sham stimulation the anode or cathode were randomly placed at F3.

For experiments 2-5, all participants underwent anode and sham stimulation in separate sessions. An additional session of cathodal stimulation was added to experiments 1 and 6 (Table 5.1). The order of the sessions was randomised and counterbalanced across participants. To minimise any potential carry-over effect, there was at least 48hr gap between each stimulation session (Nitsche et al., 2008b). Stimulation in each session lasted 25 minutes, including 30 s to ramp-up and down the stimulating current. For the sham condition, electrical current was only applied during the first and last 30 s of the stimulation, so as to induce the same cutaneous sensation as real stimulation, and thus blind the participants as to stimulation condition.

During the first 5 minutes of stimulation, participants sat relaxed with eyes closed. This delay was designed to allow potential neuro-modulatory effects to build up (Zwissler et al., 2014). Next, participants began the behavioural task while stimulation continued. All participants finished the behavioural task within around 20 min, coincident with the end of stimulation. In case participants finished the task prior to the end of stimulation they were asked to remain seated until the end of the stimulation. In case the task outlasted the stimulation, they continued to perform the task without further stimulation. The task period never exceeded the stimulation period by more than 2 min. No adverse effect of stimulation was reported by the participants other than mild tingling sensation.

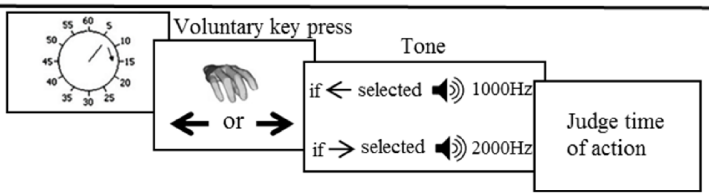
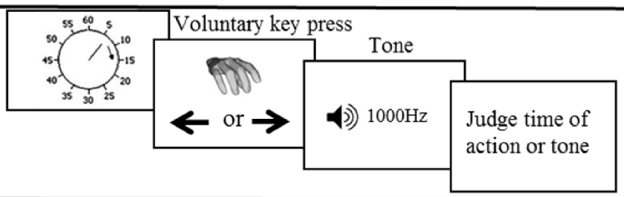
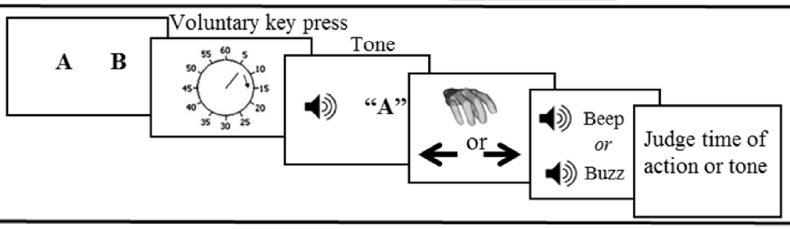
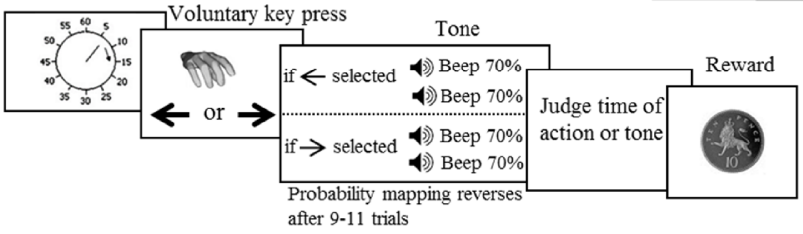
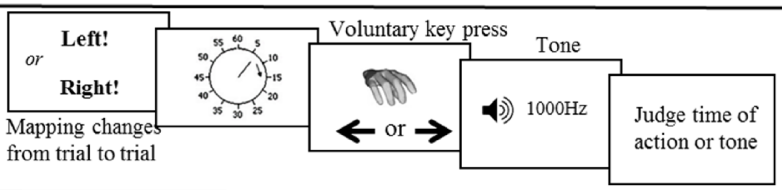
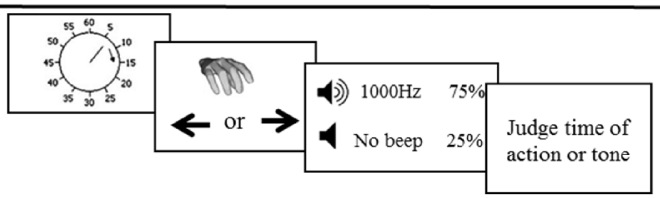
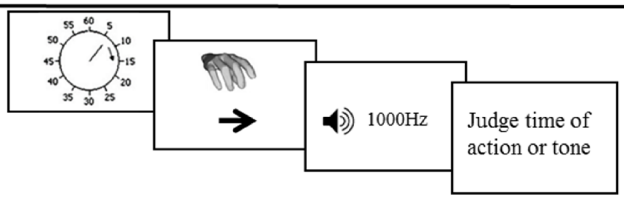
Experiment	Design	tDCS montage	Dependent variable
Exp. 1 *Endogenously selected actions* *Different outcome identity*		Anode Cathode Sham	Action binding
Exp. 2 *Endogenously selected actions* *Same outcome identity*		Anode Sham	Action binding Tone binding
Exp. 3 *Rule-based endogenously selected actions* *Different outcome identity*		Anode Sham	Action binding Tone binding
Exp. 4 *Endogenously selected actions* *Different outcome value*		Anode Sham	Action binding Tone binding
Exp. 5 *Exogenously selected actions* *Same outcome identity*		Anode Sham	Action binding Tone binding
Exp. 6 *Endogenously selected actions* *Uncertain outcome*		Anode Cathode Sham	Action binding
Published Dataset *Fixed action* *Fixed outcome*		Anode Sham	Action binding Tone binding

Table 5.1. Experimental designs, tDCS montage and dependent variables of Experiments 1-6 and dataset 7.

5.2.4. Data analysis

For each condition a judgement error was calculated as the difference between judged and actual clock time. The averaged judgement error across trials was then calculated for each condition. 'Action binding' was defined as the difference between the mean judgement error in the operant action and the baseline action conditions. A positive value represents perceptual shift of the action toward its outcome. 'Tone binding' was defined as the difference between the judgement error in operant tone and baseline tone conditions. A negative value of tone binding represents the perceptual shift of outcome toward its action.

One-way repeated measure ANOVA and paired-samples t-test were used for comparisons within each experiment (anode and/or cathode vs. sham).

We used two statistical methods to investigate commonality across experiments. First, we performed meta-analyses using both fixed effects and random effects models (Borenstein et al., 2011). The analysis followed the steps described by Lipsey & Wilson (Lipsey and Wilson, 2001), and made use of freely-available software for plotting (Gordon and Lumley, 2015). Effect sizes were calculated using Cohen's d_{av}, with Hedges' correction for sample size biases (Hedges's g_{av}) (Lakens, 2013). This measure is recommended for within-subjects designs, because of its ready compatibility with other more familiar between-subject effect size measures, and because alternative effect size measures such as g_{rm} can be overconservative when observations are highly correlated across conditions (Lakens, 2013). To investigate the heterogeneity between our experiments we both calculated the Cochran's *Q* test statistic, and we also quantified degree of inconsistency using the established I^2 measure (Higgins et al., 2003).

Finally, we also pooled data across experiments, and used a mixed design ANOVA with the within subject factor of stimulation and the between subject fixed-effect factor of experiment.

5.3. Results

5.3.1. Experiment 1: endogenously selected actions with different outcome identities

18 participants, 18-32 years of age (mean 23, 10 females) were recruited (data from two were unavailable due to technical errors). They performed the 'intentional binding' task with the following modifications: participants freely chose an action between a pair of action alternatives (left or right arrow). Left keypress produced a low pitched beep (1000Hz, 100ms) and right keypress produced a high pitched beep (2000Hz, 100ms) after 250ms. Participants then made a judgement about the time of their keypress (action binding). To ensure attention to tones, beeps were sometimes associated with frequency modulation. Participants were asked at the end of each block to rank the order of the beeps according to how often they heard them. All participants were exposed to anode, cathode and sham stimulation of left DLPFC in different sessions.

One-way repeated measure ANOVA with the factor of stimulation type showed that action binding was not significantly modulated by the stimulation ($F(2,30) = 3.12$, $p = 0.06$, $\eta 2 = 0.17$).

5.3.2. Experiment 2: endogenously selected actions with the same outcome identity

16 participants, 18-28 years of age (mean 20, 8 females) were recruited. They were instructed to freely choose between two action alternatives (left or right arrow). Either keypress produced the same outcome (1000Hz, 100ms tone after 250ms). Participants judged the time of their keypress (action binding) or the beep (tone binding) in separate blocks. All participants were exposed to anode and sham stimulation of left DLPFC in different sessions.

Paired-samples t-test showed a significant increase in action binding after anodal stimulation of the left DLPFC compared to sham ($t(15) = 2.55$, $p = 0.02$, $g_{av} = 0.81$). No significant effect of stimulation was observed on tone binding ($t(15) = -0.61$, $p = 0.55$, $g_{av} = -0.06$).

5.3.3. Experiment 3: rule-based endogenously selected actions with different outcome identities

16 participants, 18-30 years of age (mean 22, 10 females) were recruited. They had to remember the spatial location of letters 'A' and 'B', which were presented on the screen at the start of each trial. Then the computer presented vocal instructions 'A' or 'B' at a random time after the clock started. Participants thus maintained an auditory-motor mapping in working memory, and used it to select the left or right response, and execute this when they wished to (e.g., if letter 'A' was on the left and 'B' on the right side, it means respond left if you hear 'A' and right if you hear 'B') (Rowe et al., 2000). Actions were followed 250 ms by a beep (2000Hz, 100 ms) for correct responses and a buzz (500Hz, 100 ms) for incorrect. Participants were instructed to report the time of their keypress or tone in different blocks. The mapping changed randomly from trial to trial. All participants were exposed to anode and sham stimulation of left DLPFC in different sessions.

There was a significant increase in action binding after anodal stimulation of the left DLPFC compared to sham ($t(15) = 2.21$, $p = 0.04$, $g_{av} = 0.95$). No significant effect of stimulation was observed on tone binding ($t(15) = -0.35$, $p = 0.73$, $g_{av} = -0.09$).

5.3.4. Experiment 4: endogenously selected actions with different outcome values

16 participants, 18-26 years of age (mean 20, 9 females) were recruited. They freely chose to press either the left or right arrow keys. One key caused a beep (1000Hz, 100 ms) with probability of 0.7 and the other key with probability of 0.3 after 250 ms. On the remaining trials, a buzz (500Hz, 100 ms) was played instead of the beep. The probability mapping of keys to tones reversed after a run of 9-11 trials (randomised). Beeps were associated with monetary reward (£0.10). Buzzes were not rewarded. Participants judged the time of their keypress or the tone in separate blocks. All participants were exposed to anode and sham stimulation of left DLPFC in different sessions.

We found no significant effect of stimulation on action (t(15) = 1.11, p = 0.28, g_{av} = 0.19) or tone binding (t(15) = 0.38, p = 0.71, g_{av} = 0.04).

5.3.5. Experiment 5: exogenously selected actions with the same outcome

16 participants, 18-27 years of age (mean 23, 10 females) were recruited. They were explicitly instructed to press the left or right arrow keys (randomised) at the beginning of each trial at a time of their own free choice. Both actions resulted in a 1000 Hz, 100 ms tone after 250 ms. There was no 'correct' action, nor financial reward for making one action rather than the other. Participants judged the time of their keypress or the tone in separate blocks. All participants were exposed to anode and sham stimulation of left DLPFC in different sessions.

We found no significant effect of stimulation on action (t(15) = 0.48, p = 0.63, g_{av} = 0.16) or tone binding (t(15) = 0.55, p = 0.59, g_{av} = 0.15).

5.3.6. Experiment 6: endogenously selected actions with uncertain outcomes

18 participants, 18-26 years of age (mean 21, 12 females) were recruited. They freely chose between left or right arrow keypresses. Each keypress produced a beep (1000Hz, 100 ms) after 250 ms with a probability of 0.7. No tone occurred on the remaining trials. Participants then made judgements about the time of their keypress. All participants were exposed to anode, cathode and sham stimulation of left DLPFC in separate sessions.

One-way repeated measure ANOVA with the factor of stimulation type showed no significant effect of stimulation on action binding (F(2,34) = 0.08, p = 0.92, η2 = 0.01).

5.3.7. Dataset 7: fixed responding with no action alternative and fixed outcome

We previously showed that anodal stimulation of left DLPFC in a task involving no action selection affected intentional binding compared to sham (chapter 4). Our original analysis used linear discriminants to identify the combination of action and tone binding that

maximally distinguished these two conditions (mean(se) anode action binding = 25(10) ms, mean(se) sham action binding = 52(21) ms, t(15) = -1.26, p = 0.22, g_{av} = -0.42; mean(se) anode tone binding = -56(21) ms, mean(se) sham tone binding = -99(25) ms, t(15) = 2.07, p = 0.06, g_{av} = 0.45).

5.3.8. Meta-analysis:

The overall picture emerging from individual experiments is complex, with some experiments producing a significant result, and others not. The individual results are summarised in Table 5.2. The experiments were not direct replications, since the behavioural paradigms and stimulation conditions included varied between experiments. However, anodal and sham stimulation of the left DLPFC were common conditions, and action binding was a common dependent variable, in all seven experiments. We did not perform any other experiments involving these conditions and measures. Therefore, we had a large dataset, free from publication bias, in which we could investigate the following series of research questions:

1. Is there an overall effect of DLPFC tDCS on action binding?
2. Is there heterogeneity between our experiments that cannot be explained by chance alone?
3. If there is heterogeneity, is there a plausible classification of the experiments into subsets of studies that can explain the heterogeneity?
4. Is there a significant effect of frontal tDCS within each such subset?

Meta-analysis of laboratory interventions are rare (Horvath et al., 2014, 2015a), and have received less statistical attention than clinical trials – a point to which we will return in the discussion.

Two methodological points about our meta-analysis require specific mention. First, we performed both fixed-effects and random-effects analyses. Whether fixed or random-effects analyses are more appropriate remains controversial, and has not been systematically explored for laboratory experiments, particularly for cases where heterogeneity between

subsets of studies may be present (Hedges and Vevea, 1998; Horvath et al., 2015a). We return to this point in discussion.

Experiment	Results	
Exp. 1 *Endogenously selected actions Different outcome identity*	action binding was not significantly modulated by the stimulation.	F(2,30) = 3.12, p = 0.06, η2 = 0.17
Exp. 2 *Endogenously selected actions Same outcome identity*	significant increase in action binding after anodal stimulation of the left DLPFC compared to sham	t(15) = 2.55, p = 0.02, g_{av} = 0.81
	No significant effect of stimulation was observed on tone binding	t(15) = -0.61, p = 0.55, g_{av} = -0.06
Exp. 3 *Rule-based endogenously selected actions Different outcome identity*	significant increase in action binding after anodal stimulation of the left DLPFC compared to sham	t(15) = 2.21, p = 0.04, g_{av} = 0.95
	No significant effect of stimulation was observed on tone binding	(t(15) = -0.35, p = 0.73, g_{av} = -0.09)
Exp. 4 *Endogenously selected actions Different outcome value*	no significant effect of stimulation on action binding	(t(15) = 1.11, p = 0.28, g_{av} = 0.19)
	no significant effect of stimulation on tone binding	(t(15) = 0.38, p = 0.71, g_{av} = 0.04).
Exp. 5 *Exogenously selected actions Same outcome identity*	no significant effect of stimulation on action binding	(t(15) = 0.48, p = 0.63, g_{av} = 0.16)
	no significant effect of stimulation on tone binding	t(15) = 0.55, p = 0.59, g_{av} = 0.15
Exp. 6 *Endogenously selected actions Uncertain outcome*	no significant effect of stimulation on action binding	F(2,34) = 0.08, p = 0.92, η2 = 0.01
Published Dataset *Fixed action Fixed outcome*	anode action binding = 25 ms, sham action binding = 52 ms, g_{av} = -0.42; anode tone binding = -56 ms, sham tone binding = -99 ms, g_{av} = 0.45	

Table 5.2. Experimental results of Experiments 1-6 and dataset 7.

Second, meta-analyses are much more common for between-subject designs than for within-subject designs, reflecting the traditional association between meta-analysis and

clinical trials. The effect size for a within-subject design can be calculated from the difference scores between the two conditions of interest, whereas the effect size for between-subject designs is based on estimates of variability in each of the two conditions. Meta-analyses of within-subject studies often report effect sizes based on between-subject effect sizes (Lakens, 2013; Horvath et al., 2015a), for comparability with other meta-analyses, and because information about variability of difference scores is rarely given explicitly in published reports of within-subjects designs. If the variability in each condition of a within-subjects design is largely due to factors common to all conditions, such as individual differences in personality, task performance etc., then the variability of difference scores may be substantially lower than the pooled variance of each experimental condition. In that case, using a between-subject method of calculating error variance would underestimate the true within-subject effect size. We have used a measure of effect size that has been recommended for within-subjects designs, and uses the average standard deviation of both conditions (Lakens, 2013).

The results of the meta-analysis are summarised in table 5.3. Regarding our first question of overall effect size, we found a near-significant overall effect of frontal tDCS when using fixed effects, but not random effects analysis. Thus, we found a trend for frontal tDCS to influence our measure of sense of agency.

Regarding our second research question of heterogeneity, we used the established I^2 measure to estimate the variability among effect sizes that could not be explained by chance. This amounted to 44% (moderate heterogeneity (Higgins et al., 2003)) under the fixed effects model. Because we purposely designed the experiments with different behavioural paradigms, some heterogeneity is expected a priori. Nevertheless, we also applied the conventional Cochrane's Q test, and confirmed presence of heterogeneity under the fixed effects model, but not the random effects model. Note that we used the recommended significance level of 0.1 for the Q test, rather than the conventional 0.05, given the acknowledgedly conservative nature of the test (Higgins and Green, n.d.).

Where heterogeneity exists, the Q test can also be used to identify potential subsets of studies that have a common pattern of results, since heterogeneity should be high between subsets, but low within. Based on previous literature (Rowe et al., 2000), we hypothesised that DLPFC stimulation might influence action binding differently when participants themselves select between alternative actions (experiments 1, 2, 3, 4, 6), compared to when they did not (experiment 5, dataset 7). This subsetting of studies explained a high (71%) proportion of the variability between experiments, and revealed a significant difference between experiments with and without endogenous action selection (Table 5.3 & Fig. 5.1). Importantly, within the action selection subset, we found a highly significant effect of stimulation (mean (se) g_{av} = 0.41(0.16), z-test = 2.56, p = 0.01, 95% CI = [0.10 0.72]) with low and non-significant heterogeneity (I^2_{Subset}= 31%, $Q_{SubsetTotal}$(4) = 5.79, p > 0.1). In the subset of experiments lacking action selection, we found no effect of stimulation (mean (se) g_{av} = -0.14(0.25), z-test = -0.57, p = 0.57, 95% CI = [-0.63 0.35]).

Research Question	***Analysis Model***	***Effect size/ Test statistic***
Is there an overall effect of stimulation?	Fixed effects	mean (se) gav = 0.25 (0.13) z-test = 1.84, p = 0.06 95% CI = [-0.02 0.51]
	Random effects	mean (se) gav = 0.26 (0.18) z-test = 1.45, p = 0.15 95% CI = [-0.09 0.61]
Is there heterogeneity among experiments?	Fixed effects	***I^2* = 44% (moderate heterogeneity)** ***Q_{Total}* (6) = 10.66, p < 0.1**
	Random effects	I^2= 1% (negligible heterogeneity) Q_{Total}(6) = 6.05, p > 0.1
Is there a convincing subset of studies where an effect is present? ***Grouping based on action selection***	Fixed effects	***$I^2_{Between}$= 71% (high heterogeneity)** ***$Q_{Between}$* (1) = 3.47, p < 0.1**
	Random effects	$I^2_{Between}$ = 50% (moderate heterogeneity) $Q_{Between}$(1) = 2.00, p > 0.1

Table 5.3. Meta-analysis of the experimental results, presented separately for fixed and random effect models. Significant results are presented in bold.

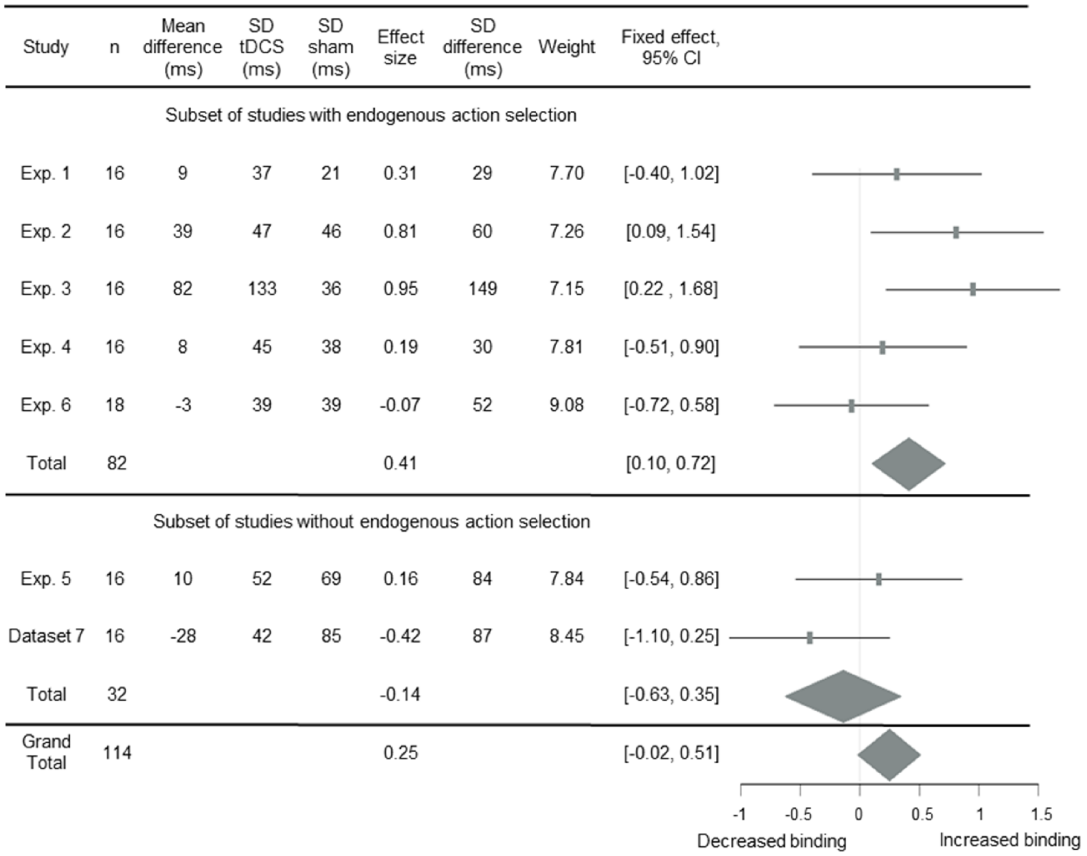


Figure 5.1. Forest plot comparing subsets of studies with and without endogenous action selection.

Unusually for a meta-analysis, we had ready access to the original data, not just effect sizes. We therefore also conducted a mixed effects ANOVA across the 114 participants, using a within-subject factor of the stimulation (anode vs. sham) and a between-subject factor of experiment (7 levels, Exp1-6 + dataset 7), with experiment as a fixed effect. We found no significant main effect of experiment ($F(6,107) = 0.88$, $p = 0.51$, $\eta 2 = 0.05$), but a significant overall increase in action binding with stimulation ($F(1,107) = 5.13$, $p = 0.03$, $\eta 2 = 0.05$), and a significant interaction between the stimulation and experiment ($F(6,107) = 3.09$, $p < 0.01$, $\eta 2 = 0.15$) (Fig.5.2).

We conducted a further ANOVA, collecting the experiments into the subsets of studies identified by meta-analysis. This showed no significant main effect of stimulation ($F(1,112) = 1.06$, $p = 0.30$, $\eta 2 = 0.01$) nor of subsets ($F(1,112) = 0.48$, $p = 0.50$, $\eta 2 < 0.01$), but a significant interaction between the stimulation and the subsets ($F(1,112) = 4.10$, $p = 0.04$, $\eta 2 = 0.04$). Anodal stimulation increased action binding within the endogenous action-selection subset of studies ($t(81) = 2.93$, $p < 0.01$, $g_{av} = 0.47$), but not in the other subset ($t(31) = -0.56$, $p = 0.58$, $g_{av} = -0.13$).

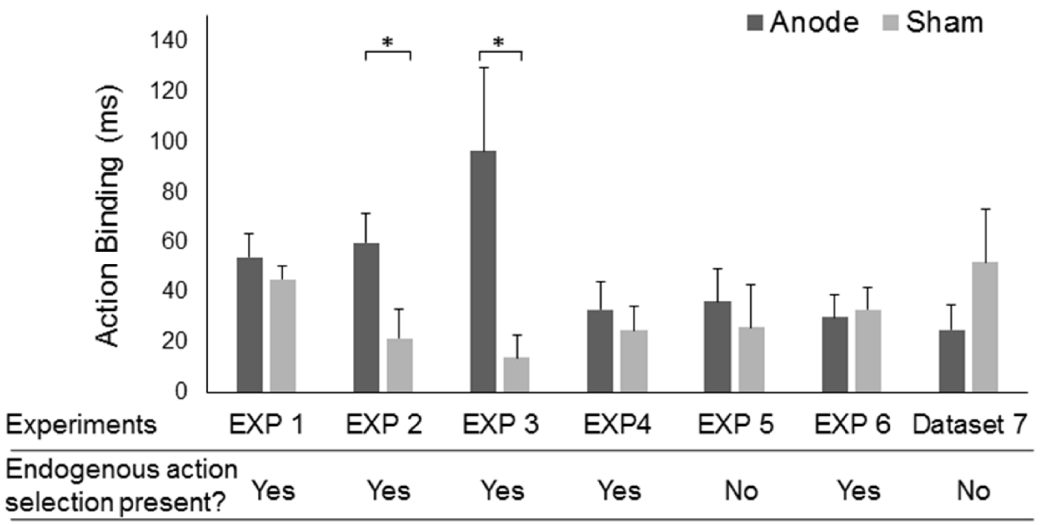


Figure 5.2. The effect of anodal stimulation of the left DLPFC on action binding in each of seven studies. Error bars show standard error of the mean. *$p < .05$.

5.4. Discussion

5.4.1. Effect of anodal stimulation of DLPFC on action binding: a meta-analytic approach

DLPFC is routinely activated in studies of human voluntary action (Jahanshahi et al., 1995), particularly when a strong action selection component is present (Rowe et al., 2000). Moreover, a recent study involving explicit agency judgement found that when action selection was facilitated by compatible primes, the DLPFC influenced sense of agency over

outcomes, through its connectivity with angular gyrus (Chambon et al., 2013). Our previous neurostimulation experiment found some evidence in support of the involvement of lateral frontal areas in sense of agency (Chapter 4; Khalighinejad and Haggard, 2015). However, that task lacked any action selection component. To further explore the contribution of DLPFC to sense of agency, we performed a series of tDCS experiments, altering task parameters related to action selection and action outcome, and measuring the temporal association between a voluntary action and its outcome, as an implicit measure of sense of agency.

In a series of experiments, participants were asked to endogenously select an action from two action alternatives. Actions were then followed by a sensory outcome (tone) of different identities (Exp. 1, 3), same identity (Exp. 2), different outcome values (Exp. 4), or uncertain outcomes (Exp. 6). In other experiments, participants were explicitly told which action to choose (Exp. 5) or had no alternatives (dataset 7). Primary analysis of individual experiments showed that anodal stimulation of the left DLPFC sometimes increased action binding but sometimes did not. We then used meta-analysis to investigate the results across the several experiments.

Fixed effects meta-analysis confirmed moderate heterogeneity among studies. Classifying studies into subsets, according to presence or absence of endogenous action selection, explained 71% of the variability across the experiments. Within the subset of experiments involving action selection, anodal tDCS significantly increased action binding compared to sham. We conclude that anodal stimulation may have a modest (mean g_{av} = 0.41) facilitatory effect on a component of intentional binding related to selection between alternative voluntary actions. Thus, our meta-analysis suggests a causal role of frontal action selection signals in prospective sense of agency. A similar suggestion was made previously in the context of a neuroimaging study using explicit agency judgements (Chambon et al., 2013). Our result provides the first evidence using neurostimulation, and implicit measures.

5.4.2. DLPFC, action selection, and sense of agency

Few causal studies have investigated the neural correlates of agency, and those focussed primarily on medial frontal cortex. Direct stimulation of medial frontal cortex in neurosurgical patients can induce an experience of volition ("urge to move") (Fried et al., 1991). In contrast, stimulation of lateral frontal cortex produced movements without any subjective experience of urge (Desmurget et al., 2009), though the stimulation sites were largely posterior to DLPFC.

Three studies have reported neurostimulation effects on sense of agency over outcomes. One study combined a choice reaction-task with single pulse TMS stimulation of the inferior parietal cortex (IPC), or the DLPFC, or sham stimulation (Chambon et al., 2014a). Participants made explicit judgements of control over action outcomes. There was no main effect of TMS on judgements of control, but stimulation of IPC at the time of action selection reduced the influence on judgements of compatible vs. incompatible subliminal primes preceding the instructional cue. The absence of any significant effect of DLPFC TMS may seem inconsistent with the present results. However, in that task, the action selection was direct and exogenous. In contrast, in our experiments, action selection always involved some endogenous element, such as learning to identify a 'correct' response, or an indirect cue-action mapping requiring a working memory component. Two other studies focussed on medial frontal areas. A cTBS virtual lesion study of pre-supplementary motor area (pre-SMA) was shown to reduce effect binding, but not action binding, in a task where participants could make only a single action, but chose the time at which to do so (Moore et al., 2010). Finally, a recent tDCS experiment (Cavazzana et al., 2015) combined intentional binding with stimulation of pre-SMA. Those authors found reduced action binding, but not effect binding, again in a task that lacked action selection.

5.4.3. Limitations

We always placed the return (cathode) electrode on the right supraorbital area. Although this montage is common in tDCS studies of prefrontal areas (Tremblay et al., 2014), it involves a high current density close to the right frontopolar area. Hence, we cannot completely rule out effects of right cathodal frontopolar stimulation. However, we think this alternative interpretation is unlikely for a number of reasons. First, a recent review found that cathodal stimulation has only weak effects on cognitive (as opposed to motor) measures (Jacobson et al., 2012b). Second, in previous experiments focussing on the parietal contribution to sense of agency, we found strong effects of parietal anodal stimulation, both with cathode placed over the right supraorbital area, or placed elsewhere (Chapter 4; Khalighinejad and Haggard, 2015). Third, an overview of studies of anodal left DLPFC studies shows effects both with supraorbital and other cathodal locations, including left mastoid (Zaehle et al., 2011) and Cz (Meiron and Lavidor, 2013). Thus, while we cannot strictly rule out a right frontopolar cathodal effect, we favour an interpretation based on a left DLPFC anodal effect.

Meta-analytic methods originated in the context of clinical trials, and do not transfer straightforwardly to experimental settings. In the 'ideal' clinical trial case, meta-analysis would include several large studies all using a common intervention, on comparable populations, and with a single outcome measure. In contrast, the neurostimulation literature typically provides many low-powered studies. Variations in equipment, electrode placement, and experimenter technique may influence effect size (Nitsche et al., 2015). Finally, the neurostimulation literature is probably strongly affected by publication bias.

Fixed effects models assume that all studies share a common true effect size, while random effects models assume that the true effect sizes may vary across studies (Borenstein et al., 2011). Previous meta-analyses of cognitive tDCS effects argued in favour of fixed-effects models (Horvath et al., 2015b), on the grounds that the intervention was a uniform and precisely-quantified intervention. In contrast, random-effects meta-analyses are often used in

systematic reviews of clinical trials to address likely but uncontrolled variation across trials, including factors such as differences in protocol, trial setting, etc. In our case, we varied the details of the cognitive task, and therefore might *expect* our intervention to have varying effects. However, our experimental designs followed the classic logic of cognitive neuropsychology, in which a task involves a particular combination of cognitive processes, each implemented by a specific functional circuit within the brain, which is targeted by tDCS. Finally, combining a fixed-effects model with heterogeneity analysis may seem paradoxical, since the former assumes a common effect, while the latter looks for variability. However, since the meta-analysis aims to identifying possible subsets of studies, the assumption of a single true effect size would be warranted *within each subset*. Here, we identified a subset of agency tasks involving endogenous selection between alternatives, for which DLPFC stimulation increased sense of agency. This result fits well with the neurocognitive theory of localisation of function (Shallice, 1988), and are consistent with the role for DLPFC in both action selection (Rowe et al., 2000) and prospective agency (Chambon et al., 2013). Other meta-analytic approaches have also been proposed to address such research questions. In particular, clinical trial and policy meta-analyses often favour random-effects meta-regression, using covariates to identify heterogeneity and possible subsets. However, these methods are not recommended when fewer than ten studies are available (Higgins and Green, n.d.).

The effect of anodal stimulation on action binding was not observed in every experiment, even within the subset of studies involving endogenous action selection. In the fixed-effects statistical model, the measured effects combine both the “true” effect, and measurement error. Measurement error may thus explain why not all individual studies showed the significant effect found in the subset of action selection studies. Indeed, estimating the true effect size, independent of measurement error, is one explicit aim of meta-analysis. Further, our studies differed in other aspects, such as outcome identity and value, in addition to the common feature of endogenous action selection. We cannot exclude the possibility that left

anodal DLPFC stimulation might also have some influence on other cognitive processes involved in some of the experiments – and this could partly explain the varied effect of stimulation on action selection subset of studies. In this vein, we also applied similar meta-analytic analyses of heterogeneity to investigate whether stimulation effects depended on the nature of action outcomes in our data. These analyses did not reveal any significant pattern in the results. However, small additional effects, related to cognitive processes other than action selection, might still be present in some individual experiments.

5.4.4. Conclusions

We performed a series of experiments to test whether we can interfere with sense of agency in the context of action selection, by combining frontal tDCS, with an implicit measure of sense of agency based on mental chronometry. Anodal stimulation of DLPFC increased binding of actions towards outcomes, but only in tasks where participants endogenously selected between alternative actions. One previous behavioural study noted a relation between action selection and intentional binding (Barlas and Obhi, 2013), but the underlying neural basis remained unclear. Our result has important implications for the sense of agency. In particular, it seems incompatible with a strongly reconstructionist view that people infer agency from the mere conjunction of action execution and sensory outcomes (Wegner, 2003b). Rather, neural processes in DLPFC of selecting *which action* to make, which necessarily precede action initiation, make a prospective contribution to sense of agency. A recent meta-analysis on the efficacy of tDCS in the treatment of depression indicated that anodal stimulation of left DLPFC was superior to sham tDCS (Meron et al., n.d.). Depression and sense of agency may be related. We speculate that clinical benefits in depression could be related to increased feeling that one's decisions and actions can make a difference - an enhanced sense of agency. Finally, we note that the size of our effect is modest, and that no established statistical analysis plan exists for meta-analysing laboratory intervention studies. Our approach was based on fixed effect meta-analysis, and used heterogeneity analysis to identify subsets of studies, as a means of identifying the specific cognitive processes in

DLPFC that contribute to sense of agency. We hope that this article will trigger future methodological developments in using meta-analysis of neurostimulation data to localise cognitive functions in the human brain.

Chapter 6

Associative mechanisms underlying experiences of voluntary and involuntary actions

It remains unclear whether sense of agency is hardwired, arising from specific signals within the brain's motor systems, or rather depends on associative learning, through repeated co-occurrence of voluntary movements and their outcomes. To distinguish these two models, we asked participants to trigger a tone by a voluntary keypress action. The voluntary action was always associated with an involuntary movement of the other hand. We then tested whether the combination of the involuntary movement and tone alone might now suffice to produce a sense of agency, even when the voluntary action was omitted. Sense of agency was measured using an implicit marker based on time perception, namely a shift in the perceived time of the outcome towards the action that caused it. Across two experiments, repeatedly pairing an involuntary movement with a voluntary action induced key temporal features of agency, with the outcome now perceived as shifted towards the involuntary movement. This shift required involuntary movements to have been previously associated with voluntary actions. We show that some key aspects of agency may be transferred from voluntary actions to involuntary movements. An internal volitional signal is required for the primary acquisition of agency, but, with repeated association, the involuntary movement in itself comes to produce some key temporal features of agency over the subsequent outcome. This finding may explain how humans can develop an enduring sense of agency in non-natural cases like brain-machine interfaces.

6.1. Introduction

In a series of brilliant experiments, Roger Sperry switched the nerves for flexion of the rat hind leg with the nerves for extension. After that, whenever the bottom of the foot was injured, the rat extended the foot instead of flexing it. Rats never learned to lift up the paw, and "no adaptive functioning of the nervous system took place" (Sperry, 1940). When the optic nerves of salamanders were cut, and the eyeball rotated 180 degrees, salamanders saw upside down for the rest of their lives (Sperry, 1943). These experiments suggested that key sensorimotor brain circuits are largely hardwired, and impervious to modification by experience.

Some recent results have linked the sense of agency to specific preparatory volitional signals in frontal (Fried et al., 2011) and/or parietal (Desmurget et al., 2009) areas, which then trigger voluntary motor commands passing through the "final common path" (Sherrington, 1906) of the primary motor cortex. Importantly, these 'volitional signals' were generated well before the occurrence of both action and outcome, and were strongly correlated with the subjective intention to move. Such theories suggest a hard-wired, Sperry-esque account of human volition.

In contrast, associative theories of agency deny the special status of internal volitional signals, and focus instead on co-occurrence of actions and outcomes. For example, in *ideomotor theories,* repeated association of actions and outcomes means that, over time, actions come to be represented primarily in terms of their anticipated outcomes or goal-states. By the same association, activation of the neural code for the goal event is then able to generate the voluntary action (Prinz, 1997). Stronger versions of this view suggest that people merely *infer* their own agency based on observing the combination of action and outcome. There is no direct mental access to the internal processes that cause our actions, and the experiences of will and agency are mere inferences, or even illusions (Wegner, 2003a).

Current computational models of motor control, such as the comparator model (Frith et al., 2000; Blakemore et al., 2002) have also been used to explain sense of agency. During action execution, efferent signals from motor areas are compared to predictions about the sensory consequences of the actions, such as feedback from a moving limb, or from some other external outcome of the action. These, contain elements of both the hard-wired and the associative frameworks. On the one hand, the sense of agency could depend necessarily on hard-wired efferent motor signal. On the other hand, the predictions generated by this signal must be based on learning an internal model from previous associations between efferent signals and their sensory consequences, consistent with the associative framework. It remains unclear to what extent human sense of agency is based on such hardwired signals or on learned associations.

These models make different predictions about the possibility of transferring agency from a voluntary action to another, co-occurring event. Mental properties commonly transfer from representation of one event to representations of another, notably in classical conditioning (Pavlov (1927), 2010), but it remains unclear whether this occurs also for sense of agency. We asked participants to trigger a tone by making a voluntary keypress action with one hand. The voluntary action was always associated with an involuntary movement of the other hand. We then tested whether the combination of involuntary movement and tone alone might suffice to produce a sense of agency over the tone, even when no voluntary action was now present. Theories based on hard-wired efferent signals predict no sense of agency in this condition, since the putative internal volitional signal for one's own voluntary actions is, by definition, absent for involuntary movements. In contrast, associative learning theories predict that repeated co-occurrence of a voluntary and an involuntary movement could produce associative transfer, so that involuntary movements could, by association, come to acquire the same sense of agency that characterises voluntary movements.

We therefore designed two between-subject experiments. 36 participants were recruited for the first experiment and were randomly assigned to the experimental and control groups

(Fig. 6.1). In the experimental group, self-paced voluntary keypress actions of the right hand triggered an immediate and physically-similar involuntary keypress movement of the left hand, imposed by a robotic arm (Phantom Premium, 3D Systems, South Carolina, USA). These movements were followed by a tone 250 ms later, in the operant condition. Participants could thus learn to associate voluntary action, involuntary movement, and tone. Such "learning" trials alternated with "test" trials containing only involuntary movements followed by tones, and no voluntary action. Sense of agency over the tone was measured using an implicit marker based on time perception. Participants judged the time of the tone using a rotating clock display. A shift in the perceived time of the tone towards the preceding action, is an established implicit marker of agency. This shift is compared to a baseline condition containing only a tone, but no action. Importantly, involuntary movements are not sufficient to cause perceptual shifts of the tone, and a volitional signal appears necessary (Haggard et al., 2002b; Cravo et al., 2009; Moore and Obhi, 2012). A further control group of participants also judged the time of the tone following an involuntary movement, but had never experienced any association between involuntary and voluntary movement. Kinematics of both hands' movements were monitored online using accelerometers.

Experiment 2 used the same design, but triggered involuntary movements by non-invasive brain stimulation. 36 participants were recruited and were randomly assigned to experimental and control groups. Self-paced voluntary actions of one hand were now paired with involuntary twitches of the other hand, caused by transcranial magnetic stimulation (TMS) over primary motor cortex. These learning trials again alternated with test trials containing only involuntary TMS-evoked twitches followed by tones. Motor evoked potentials (MEPs) were recorded from the first dorsal interosseous of the left hand. The control group also judged the time of the tone following an involuntary twitch, but they had never experienced any association between the twitch and any voluntary actions.

If sense of agency depends on a hard-wired efferent signal from the voluntary motor system, no amount of associative learning should be able to induce key temporal features of

experience of agency for involuntary movements followed by tones, because the necessary volitional signal is absent in this case. Conversely, if sense of agency is based on associative learning, and if such associations can transfer from volitional signals to other events, the repeated association between voluntary action and involuntary movement should suffice to support some key temporal features of experience of agency over a tone triggered by involuntary movement.

6.2. Materials and Methods

6.2.1. Participants

In total 72 healthy volunteers, aged 18-35 years of age, were recruited from the Institute of Cognitive Neuroscience subject data pool for separate experiments. All participants were right handed, had normal or corrected to normal vision, had no history or family history of seizure, epilepsy or any neurologic or psychiatric disorder and did not have any metallic or electronic object in the head. Participants affirmed that they had not participated in any brain stimulation experiment in the last 48 h, nor had consumed alcohol in the last 24 h. Participants were paid an institution-approved amount for participating in each session of the experiment.

6.2.2. Intentional binding

We used intentional binding paradigm as an implicit measure of agency. The task was based on previous studies (12), and was programmed in LabVIEW 2012 (Austin, Texas). Participants viewed a clock hand rotating on a computer screen which was located 60cm in front of the participants in a quiet room. The initial clock position was random. Clock rotation was initiated by participants pressing the return key on a keyboard. Each full rotation lasted 2560 ms. Participants were instructed to look at the centre of the clock. Depending on the trial, they made voluntary keypress by pressing the enter key with their right index finger or made involuntary keypress by pressing the left control key with their left index finger

(Experiment 1). Participants chose for themselves when to make the voluntary actions. After each key press, the clock hand stopped at a random location, participants made a time judgement according to condition (see later). Each experimental session consisted of two conditions, presented in separate blocks. At the beginning of each block, brief instructions for the relevant condition were displayed on the screen. In the baseline condition, participants were instructed to look at the clock but not to press any key. While the clock was rotating, a pure tone (1000 Hz, 100 ms duration) was played over the headphones (or a loudspeaker in the TMS experiment), 1750-4000 ms (at random) after the onset of the trial. Participants were then asked to judge the clock hand position at the time of the tone. In the operant condition, participants pressed the key at a time of their own choosing, or made an involuntary movement (depending on the trial). Each keypress (or movement) produced a tone after 250 ms. Participants had to judge the clock hand position at the time of the tone. Baseline condition was tested in two separate blocks of 15 trials, at the beginning and end of the experiment. Operant condition was tested in a single block of 90 trials between the two baseline blocks.

6.2.3. Haptic device

For Experiment 1, Phantom Premium 1.5 haptic device (3D Systems, South Carolina, USA) was used to induce involuntary movements in participants' left index finger. This high-precision device has 3 degrees of freedom and provides a range of motion approximating hand movement pivoting at the elbow. Distal phalanx of the participant's' left index finger was attached to the distal end of the device. Matlab 2014 (MathWorks, USA) was used to communicate with the device. The following specifications were used to induce a natural-looking passive keypress in the finger: Force direction = X:0/ Y:-1/ Z:0, force amplitude = 0.7 N, duration of downward movement = 200 ms (30 ms to taper), duration of upward movement = 200 ms (30 ms to taper). To block out the noise of the device at the time of force induction, the main body was shielded in a soundproof box and only the arm was left out through a small hole. To control for the similarity of the movements across the fingers,

two accelerometers were mounted on the left and right index fingers of the participants. The kinematics of the movement were monitored by the experimenter. We measured the exact time interval between the beginning of the voluntary action and involuntary movement in the learning trials by using data from accelerometers placed on the index fingers of the left and right hand. Analysis of 10 trials selected at random showed that it took 34 ms (± 4 ms sd) between the software command being sent to the robot and the left finger actually moving, due to the mechanical delays in the robot. Importantly, these delays are present in both learning and test trials, and in both the experimental and the control groups. The only difference between trial types is the use of a voluntary keypress to initiate the software command to the robot in the voluntary trials of the experimental group - the delay between the depression of the key and the initiation of the software command was 2 ms, and was consistent across 10 trials selected for analysis.

6.2.4. TMS and MEP measurement

For Experiment 2, transcranial magnetic stimulation was delivered with a Magstim 200 stimulator (Whitland, UK). The optimal location for producing twitches (Motor evoked potentials (MEPs)) in the left first dorsal interosseous (1DI) was located by systematically exploring a 1-cm grid over the hand area of the right motor cortex. The motor threshold was calculated for each subject by reducing stimulator output in 5% steps to find the lowest level at which 3 MEPs exceeding 50 μV peak amplitude were obtained from 5 successive stimulations of the relaxed 1DI. Thresholds ranged from 35% to 60% of stimulator output (mean 45%). TMS output in the experiment was set at 120% of relaxed threshold. EMG was measured from the 1DI of the left and right hand with bipolar recording from surface Ag/AgCl electrodes. These data were amplified and digitized at 2 kHz (CED 1902, Cambridge, UK).

6.2.5. Experiment procedure

Upon arriving, participants were asked to read the information sheet and fill in the consent form. In Experiment 1, participants' left index finger was attached to the distal end of the

robotic arm and was placed on the left control key. The right index finger was placed on the enter key. The intentional binding task was explained for the participants and they were familiarised with the robotic arm-induced passive movements. Two accelerometers were mounted on the left and right index fingers and participants were asked to wear headphones (Sennheiser, Germany). The experiment started with the baseline block (15 trials). In this block, participants were instructed to look at a rotating clock but not to press any key. In each trial, a tone was played and participants judged the clock hand position at the time of the tone. This block was followed by the operant block, where the tone was always caused by participants' keypress at a time of their own free choice, 250 ms later. Like the previous block, participants were asked to judge the clock hand position at the time of the tone. In the experimental group, in the first 30 trials of the operant block, voluntary keypresses of the right hand were paired with the involuntary keypresses of the left hand. These *learning* trials were followed by 30 *test* trials, where a command appeared on the screen and instructed participants not to make any voluntary keypress with their right hand. Meanwhile, at a random time, participants made a passive keypress with their left hand. As in previous trials, they made judgements about the time of the tone which followed their keypress. These test trials were interleaved with another 30 learning trials. Therefore, each operant block consisted of 60 learning trials and 30 test trials. In the control group, participants never made any voluntary action, therefore, their learning trials only consisted of passive keypress with the left hand. In both groups, the experiment finished by performing another baseline block.

Experiment 2 followed the same principles as in experiment 1 with the following differences: Participants were asked to place their left hand on the desk next to the keyboard. Robotic arm-induced movements of the left index finger were replaced with TMS-induced twitches. The TMS coil was optimally positioned in each subject to produce involuntary movement of left index finger, minimising contraction of more proximal muscles and muscles activating other joints. The headphones were replaced with loud speakers. Accelerometers were replaced with surface electrodes for electromyography (EMG). As in the previous

experiment, participants made judgements about the time of the beep in three separate blocks.

Judgement error was calculated by measuring the difference between the judged clock time and the actual time. The averaged judgement error across the trials was then calculated for each block. 'Tone binding' was defined as the difference between the judgement error in the operant and the baseline condition. The negative value of tone binding represents the perceptual shift of outcome toward its action. Tone binding data from the *test* trials only were used for analysis.

Experimental design and procedure were approved by the UCL research ethics committee, and followed the principles of the Declaration of Helsinki. Transcranial magnetic stimulation followed established safety procedures (Rossi et al., 2009).

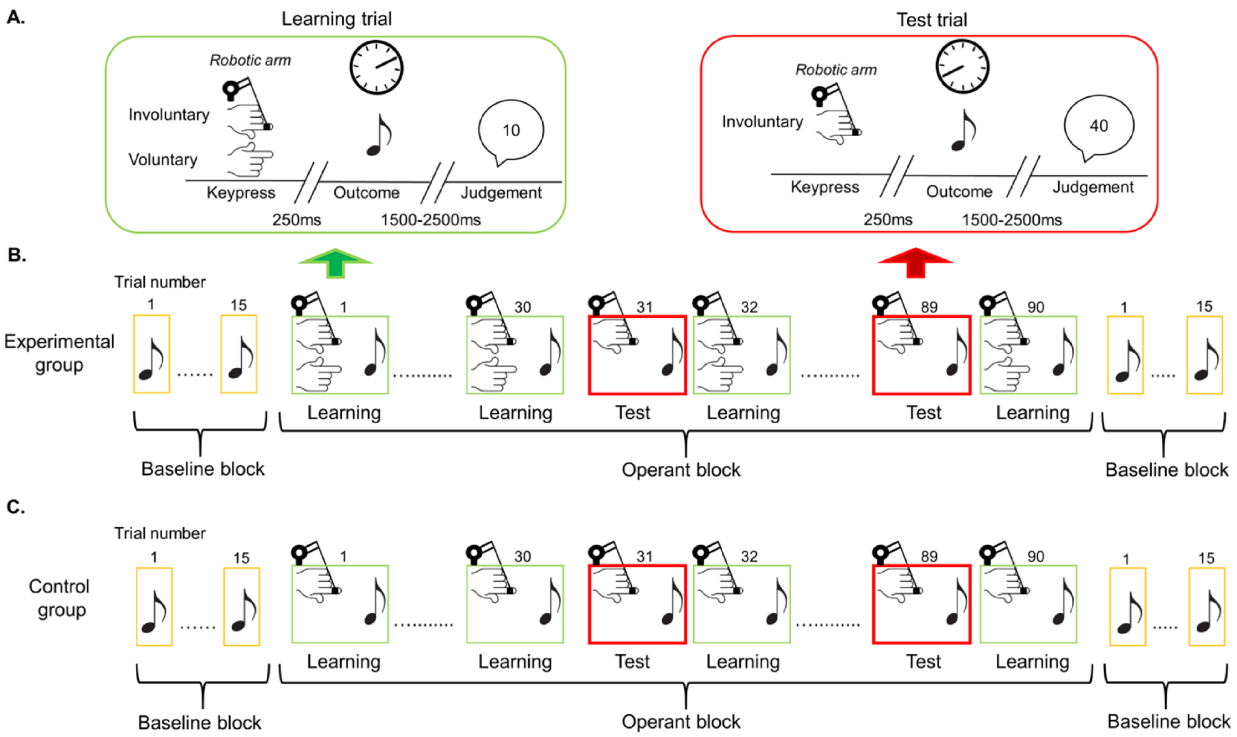


Figure 6.1.A. Timeline of an experimental trial. In the learning trials (the left green box), participants were instructed to press the enter key on a keyboard in front of them with their right index finger at a time of their own free choice. This action was paired with involuntary keypress (left control key) induced by a robotic arm pressing on the left index finger. In operant blocks, each keypress was followed by a beep (1000 Hz) after 250 ms. At the end of the trial, participants reported the perceived time of the beep. See text for full explanation. B. In the experimental group, the session started with a baseline block. The operant block then ensued. Voluntary actions of one hand were paired with involuntary movements of the other hand, followed by a tone 250 ms. After an initial learning phase of 30 trials, further learning trials were interleaved with test trials (A. right red box) on which involuntary movements were followed by tones, but no voluntary action occurred. The session ended with the execution of a further baseline block. C. A group of control participants followed the exact same design as the experimental group, but their involuntary movements were never associated with voluntary actions. In both groups, data from the test trials (red bold boxes) was used for analysis. The corresponding trial number is shown above each box. For Experiment 2, robot-induced movements were replaced with TMS-induced twitches.

6.3. Results

6.3.1. Experiment 1. Involuntary movement induced by a robotic arm

36 participants were randomly assigned to the experimental group (n=18) or the control group (n=18). Data from four participants were lost due to technical errors, leaving 16 participants in each group. We already knew, from previous evidence, that the perceived time of a tone shifts towards a preceding voluntary action, but not an involuntary movement (Haggard et al., 2002b). Surprisingly, in the experimental group, we also found a perceptual shift in the perceived time of the tone towards the involuntary movements on interleaved test trials occurring in-between truly voluntary actions (one-sample, $t(15)=-4.18$, $p<0.01$, 95% CI [-200, -65]). In the control group, who never experienced association between voluntary actions and involuntary movements, the perceived time of the beep did not shift towards the preceding movement (one-sample, $t(15)=-1.46$, $p=0.17$, 95% CI [-89, 17]). Crucially, the tone binding was significantly stronger in the test trials of the experimental group compared to the

control group (t(30) = -2.40, p=0.02, d=0.85, 95% CI [-179, -14]) (Fig. 6.2, Fig. 6.3, Fig. 6.4 and Table S.6.1).

Finally, to confirm that the difference in baseline blocks did not drive the observed effect, judgement errors from the baseline blocks were compared. No significant difference was observed between the experimental and the control group (t(30) = -0.03, p=0.97, d=0.01, 95% CI [-44, 43]).

6.3.2. Experiment 2. Involuntary movement induced by TMS

Experiment 2 aimed to replicate the first experiment and to explore any potential differences in the central and peripheral routs of passive movement induction. 36 participants were randomly assigned to the experimental group (n=18) or to the control group (n=18). Data from two participants were lost due to technical errors, leaving 17 participants per group. We replicated the core findings of experiment 1. In the experimental group, associating passive movements with voluntary actions of the other hand led to the perceptual shift of outcomes towards TMS-induced passive movements (one-sample, t(16)=-3.27, p<0.01, 95% CI [-133, -28]). No tone binding was observed in the control group (one-sample, t(16)=-0.31, p=0.76, 95% CI [-67, 50]). Direct comparison of the two groups showed a clear trend for stronger binding on test trials in the experimental group compared to the control group, though with a lower effect size that in experiment 1 (t(32) = -1.96, p=0.06, d=0.67, 95% CI [-147, 3]) (Fig.6.2, Fig.6.3, Fig.6.4, and Table S.6.2).

Finally, to confirm that the difference in baseline blocks did not drive the observed effect, judgement errors from the baseline blocks were compared. No significant difference was observed between the experimental and the control group (t(32)=0.04, p=0.97, d=0.01, 95% CI [-48, 50]).

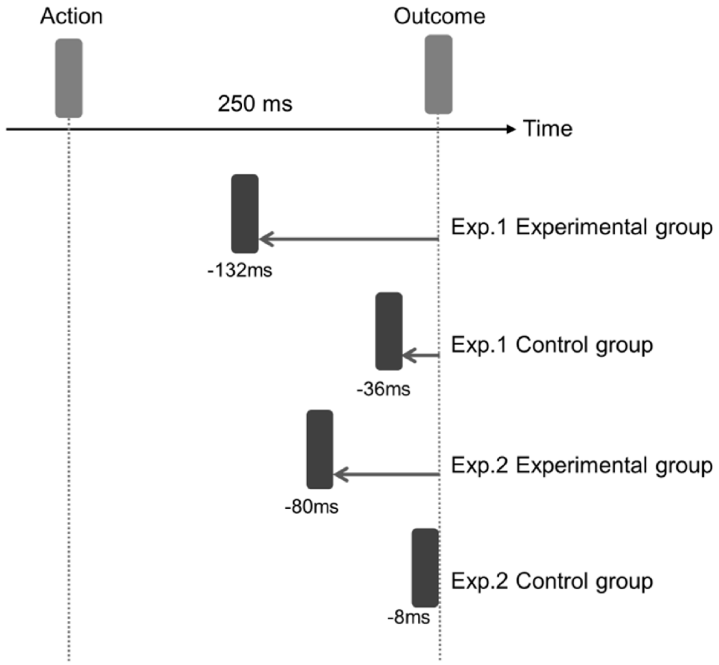
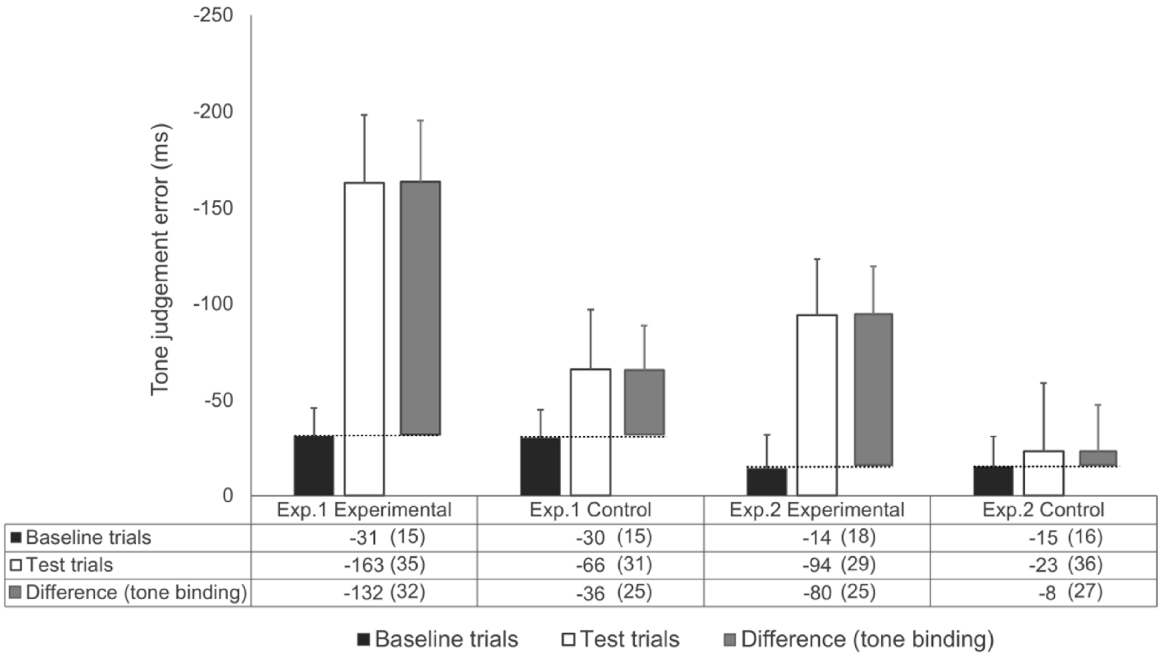


Figure 6.2. Tone binding in test trials of Experiment 1 and Experiment 2. The dashed line indicates the perceived time of the tone in the baseline condition. Binding effects are drawn to scale, and values are in ms. Differences in baseline values across sessions have been removed for display purposes.



	Exp.1 Experimental	Exp.1 Control	Exp.2 Experimental	Exp.2 Control
Baseline trials	-31 (15)	-30 (15)	-14 (18)	-15 (16)
Test trials	-163 (35)	-66 (31)	-94 (29)	-23 (36)
Difference (tone binding)	-132 (32)	-36 (25)	-80 (25)	-8 (27)

Figure 6.3. The perceived time of the tone (judgement error in ms) is shown in a baseline condition, where neither action nor involuntary movement occur, and in the involuntary movement test trials for the control and experimental groups. The difference between baseline and movement trials is an estimate of shift in the perceived time of the tone due to the preceding movement (grey bars). This "tone binding" effect serves as an implicit marker of sense of agency. Note inverted Y axis. Error bars show standard error of the mean. The mean (and standard error across participants) for all groups and experiments is presented in a table below the figure.

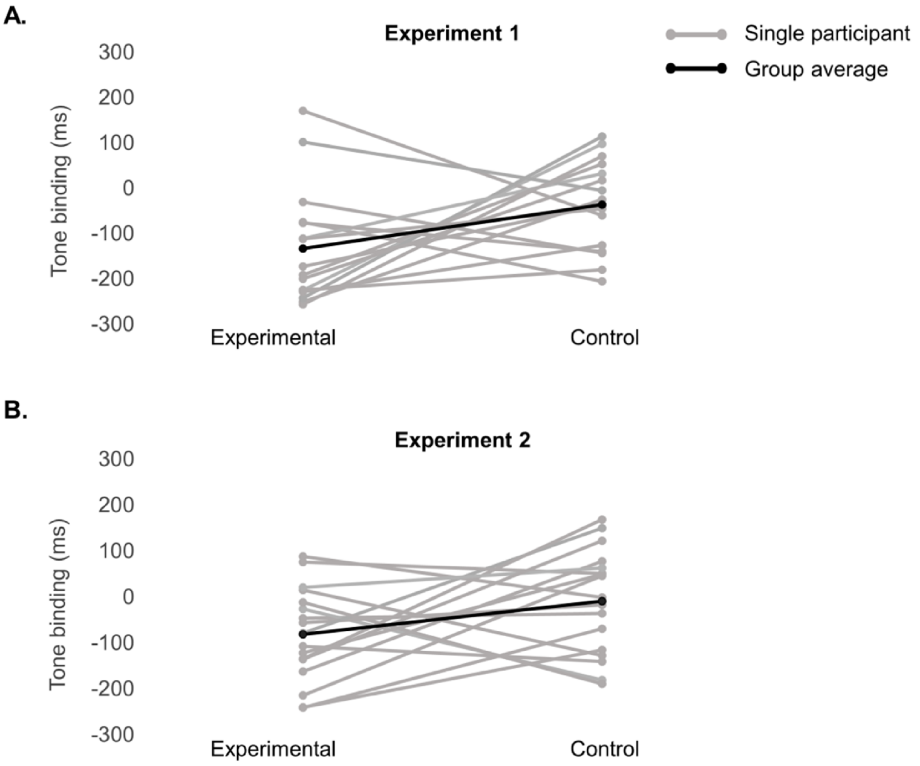


Figure 6.4. Tone binding in Experiment 1 (A) and Experiment 2 (B) for single participants in the experimental and control groups.

6.3.3. Motor evoked potentials

Any difference in amplitude of TMS-induced twitches of the left index finger between the experimental and control group could influence tone binding. To rule out any such possibility,

MEP amplitudes were compared between the two groups. No significant difference was observed between peak-to-peak amplitudes of MEPs in the test trials ($t(31) = 0.54$, $p=0.60$, $d=0.19$, 95% CI [-0.40, 0.69]) or in the learning trials ($t(31)=-0.77$, $p=0.45$, $d=0.27$, 95% CI [-0.68, 0.31]: Fig. 6.5) (MEP data from one subject was unavailable due to technical error). These results also exclude the possibility that participants could have produced some voluntary motor drive, however modest, in test trials – since even a minimal voluntary motor command, or a 'motor image', would be expected to increase corticospinal excitability (35, 36).

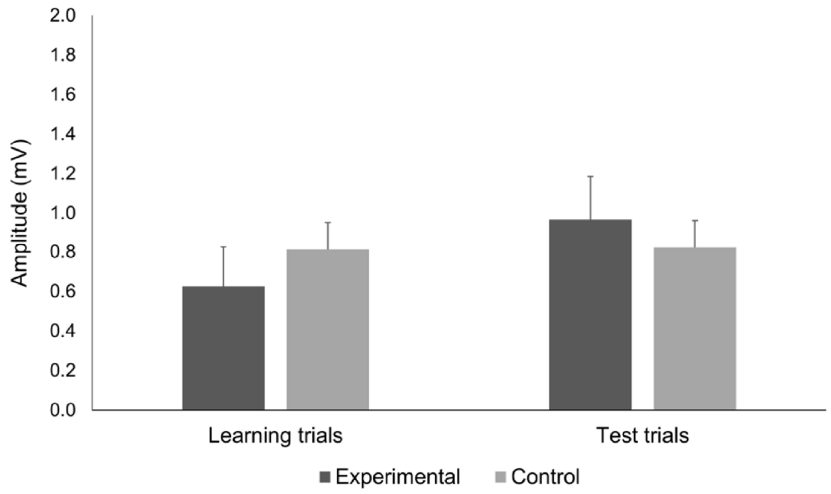


Figure 6.5. MEPs in the experimental and control groups, presented separately for the learning and test trials. No significant effects were observed. Error bars show standard error of the mean.

6.3.4. Experiment 1&2

To investigate the generality of the effect across experiments, we performed a 2x2x2 ANOVA with the within subject factor of condition (baseline vs. operant), between subject factor of experiment (Exp1 vs. Exp2) and the between subject factor of group (experimental vs. control). The significant interaction between condition and group ($F(1,62)=9.54$, $p=0.003$, $\eta 2=0.13$) recapitulated previous findings. Post-hoc analysis showed that the difference in the perceived time of the beep between two groups lays in the operant condition ($t(64)=-2.54$,

p=0.01, d=0.63, 95% CI [-150, -18]), not the baseline (t(64)=0.01, p=0.99, d<0.01, 95% CI [-32, 32]). There was no significant main effect of experiment (F(1,62)=2.67, p=0.11, η2=0.04). Importantly, there was no interaction between the condition and experiment (F(1,62)=2.12, p=0.15, η2=0.03), group and experiment (F(1,62)=0.10, p=0.75, η2<0.01) or condition, experiment and group (F(1,62)=0.20, p=0.66, η2<0.01). This suggests that the observed effect is a general phenomenon, regardless of the method used to induce passive movement (Fig. 6.2 and Fig. 6.3).

6.3.5. Standard deviation across trials

Could the stronger intentional binding in the experimental groups reflect a confounding effect of attention? For example, when subjects perform a voluntary action, they may direct attention to that action. Any other associated event might benefit from these effects, leading to stronger perceptual learning in the experimental group than in the control group who made no voluntary actions. Such attention-enhanced perceptual learning should lead to improved time estimation for the experimental group. In fact, we found an increased bias in judgment, corresponding to stronger tone binding, in the experimental group than in the control group.

In addition, we used the standard deviation of judgement errors across trials for each participant to calculate variable error, which is inversely related to the precision of temporal estimation of the tone. Improved attention to the tone would predict lower standard deviations for the voluntary group. Standard deviation of judgement errors across trials was compared in a 2x2x2 ANOVA with the within subject factor of trial type (baseline vs. test), between subject factor of group (experimental vs. control) and the between subject factor of experiment (experiment 1 vs. experiment 2). Standard deviation across trials was higher in test trials compared to baseline trials (F(1,62)=36.61, p<0.01), presumably reflecting the attentional effects of the more complex sequence of events in test conditions, particularly the co-occurrence of involuntary movement. We found no significant main effect of group (F(1,62)=0.17, p=0.68), or experiment (F(1,62)=2.01, p=0.16). Importantly, we found no

significant interaction between trial and group ($F(1,62)<0.01$, $p=0.95$), trial and experiment ($F(1,62)=0.10$, $p=0.75$), or group and experiment ($F(1,62)<0.01$, $p=0.98$), and no significant interaction between trial, group and experiment ($F(1,62)=0.33$, $p=0.57$). Thus, we found no evidence that the experimental group had improved attention to the tone on test trials (see Fig. 6.6).

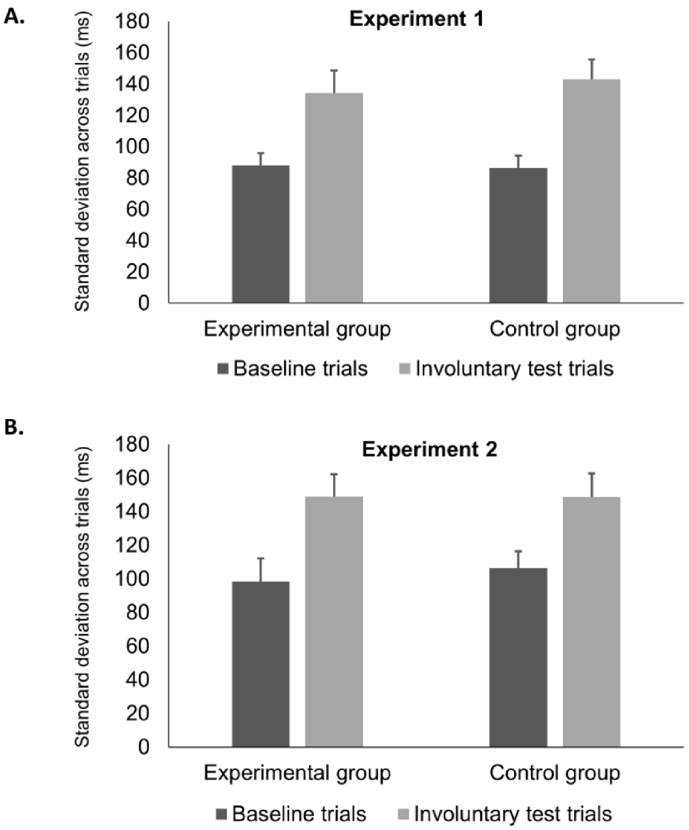


Figure 6.6. Standard deviation of tone judgement error across trials. Mean across participants of SD across trials is shown separately for baseline and involuntary movement test trials, and for experimental and control groups of experiment 1 (A) and experiment 2 (B). All values are in ms. Error bars show standard error across participants of the mean.

6.4. Discussion

'Intentional binding' refers to the perceived compression of an interval between voluntary actions and their sensory consequences. In particular, participants reliably perceive the

sensory consequences of their voluntary actions as happening earlier in time compared to a baseline condition where the same event occurs without a voluntary action. Importantly, involuntary movements were reported not to produce the same binding of tones observed after voluntary actions, but in fact produced a temporal repulsion (Haggard et al., 2002b).

Here, we showed, in two separate experiments, that repeatedly pairing an involuntary movement with a voluntary action can lead to intentional binding with respect to the involuntary movement. We have used tone binding as an implicit marker of sense of agency. Specifically, voluntary-involuntary pairing led to tone binding on involuntary movement test trials where voluntary action was absent. This acquisition of key temporal features of agency for involuntary movements did not occur in a control group who never made voluntary actions. Alternative explanations based on increased attention in the voluntary action group could not explain the pattern of results observed in our data (see supplementary results and Fig. S4). Thus, voluntary actions are necessary for the emergence of a sense of agency. However, once a voluntary signal is present, it can be mentally associated with other events, and spread to produce the distinctive intentional binding feature of volition, but now with respect to other movements. Our results therefore suggest that some key temporal features of experience of agency can be transferred by association from truly voluntary actions, to movements that are, in fact, involuntary, and purely passive.

6.4.1. A path model of agency acquisition

In our everyday life we perceive our voluntary actions as caused by our intention to produce a specific outcome. These voluntary actions are often associated with two specific experiences: The experience of volition reflects the initiation and control of the voluntary action, and possibly a prediction of the outcome. The experience of agency, in contrast, is based on attributing the actual outcome back to one's own triggering action (Haggard, 2008) (Fig.6.7.A). In our experiment, we reprogrammed the experiences surrounding voluntary action, by making participants perform two movements at the same time, one voluntary and

the other involuntary (during *learning* trials). Thus, the intention to initiate the voluntary action was associated with two movements, one located on each hand (Fig.6.7.B). Classical intentional binding predicts that experience of agency arises when there is both a direct relation between a movement and its outcome (path 2 in Fig.3.C), and also a direct relation between the movement and the intention which precedes it (path 1 in Fig.6.7.C). The necessity of path 1 is clear from previous results (Haggard et al., 2002b; Cravo et al., 2009) showing that intentional binding does not occur for involuntary movements.

In our experimental group, a further path (path 3 in Fig.6.7.C), similar to path 2, also exists between the involuntary movement and the outcome. Data from our experimental group shows that this path can generate some key temporal features of agency, such as intentional binding. Importantly, comparison with the control group shows that functioning of path 3 strongly depends on its previous association with internal volitional signals (path 1). For the control group, the involuntary movement was never paired with the voluntary action, and involuntary movements never showed the key temporal linkage to outcomes. This finding suggests that a single volitional signal can drive multiple action-outcome relations. As a result, some key temporal features of agency can arise for movements that are merely correlated with an intention, but not directly caused by it. This, in turn, suggests that the relation between intention and sense of agency is not precisely-matched, and is not effector-specific (Fig.6.7.D).

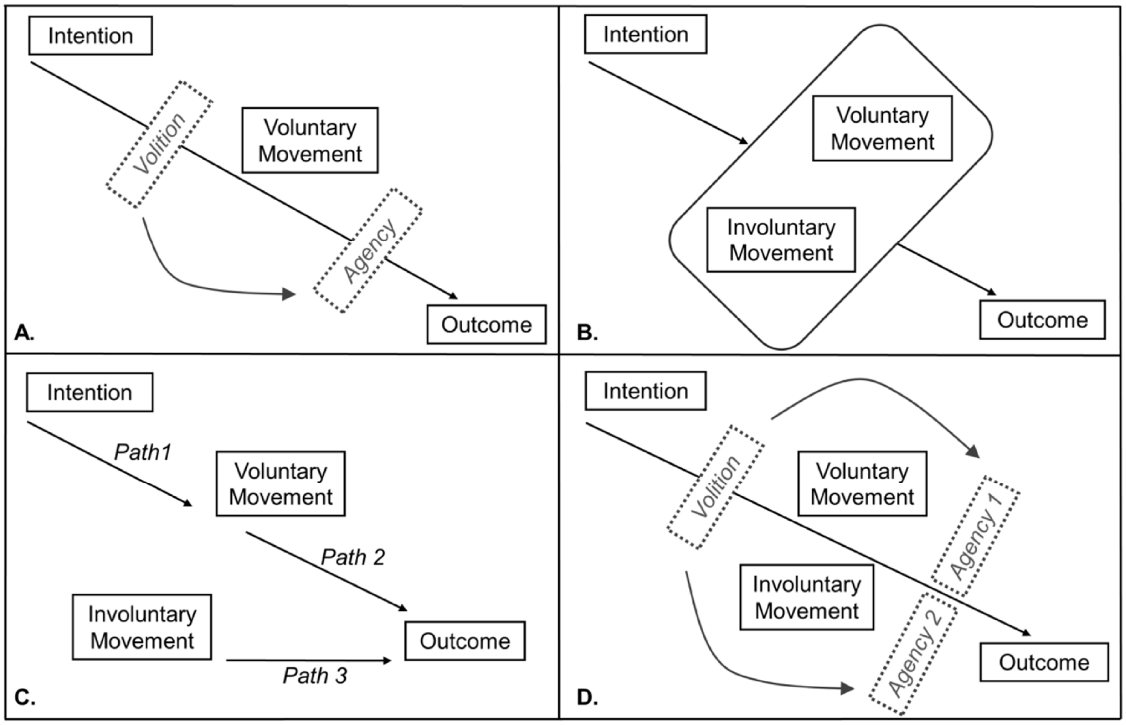


Figure 6.7. A possible mechanism for agency transfer. A. Voluntary actions are often associated with subjective experience of volition and agency. B. In our experimental group, we rewired this association so that the participants' intention produces two movements, one voluntary and the other involuntary. C. Pairing voluntary and involuntary movements lead to key temporal features of agency being experienced when involuntary movements were followed by outcomes (path 3). This experience, however, strongly depended on its previous association with intention which precedes the voluntary action (path 1). D. This suggests that once a voluntary signal is present, it can be mentally associated with other events, and spread to drive key temporal features of agency with respect to other movements.

6.4.2. Can we be mistaken regarding the facts of our own agency?

Explicit measures of agency are subject to a number of cognitive biases, and are highly sensitive to task demands. We therefore advisedly chose an implicit measure of sense of agency, based on time perception. Synofzik et al. (2008) suggested that sense of agency comprises two different levels: an implicit 'feeling of agency' and an explicit 'judgement of agency' (Synofzik et al., 2008). Based on this view, the 'feeling' of agency is produced

implicitly by low-level sensorimotor signals. In the rare case that one must explicitly judge one's own agency, this low-level feeling of agency provides the primary evidence for the judgement. However, social contextual cues and other priors can bias such judgements. The intentional binding task focusses on the non-conceptual feeling of agency. We did not obtain explicit judgements of agency in this task. As a result, we cannot know whether the involuntary movements of the experimental group came to feel like "I did that", but we imagine they did not.

In healthy adults, voluntary and involuntary movements generate quite different experiences (Wittgenstein, 2009), and our brief training was unlikely to suppress this difference. Indeed, most systems of law are based on a 'voluntary action condition', which rigidly assumes a distinct subjective experience of voluntary action (Hart and Honore, 1985). In particular, selection and preparation of action in frontal motor areas appears essential for a full experience of voluntary control (Haggard and Clark, 2003). Nevertheless, our results show that some key features of sense of agency can be transferred from voluntary to involuntary movements, given appropriate associative learning. The experience of the tone following involuntary movement acquired some temporal features of agency, but this does not imply that participants would judge themselves the author of the tone. Here, we have used implicit measures to show that one key feature of voluntary action, namely the important *'goal-directed'* or *'ideomotor'* feature, by which the experience of action leads to anticipation of outcomes, can transfer to involuntary movements. Interestingly, patients with psychosis may have a deluded experience of their own actions. These frequently involve false positives, such as reporting voluntary control over external events unrelated to their own actions, such as changing traffic lights, or news events.

Our results can also be interpreted using an active inference framework (Adams et al., 2013a). Here, intentions are abstract predictions about likely outcomes, which are Bayes-optimally combined with sensory evidence about outcomes when this becomes available. Intentional binding has been modelled as a Bayes-optimal integration of action and outcome

(Moore and Fletcher, 2012). Thus, strong tone binding might arise because intentional actions provide a high-precision prior for estimating outcomes. We found that pairing a second event, in this case an involuntary movement, with a high-precision intentional prior results in that event having a similar influence on outcome perception to the original intentional action. Thus, the structuring effects of voluntary action on outcome perception may not reflect some unique experiential quality specific to volition (though see ref. (Fried et al., 1991)), but simply that intentional actions normally serve as high-precision priors for their outcomes.

6.4.3. Specificity of Internal volitional signals underlying agency acquisition

We conclude that an internal volitional signal is required for the acquisition of sense of agency. However, after repeated association, the volitional signal is not required for subsequent expression of key temporal features of agency, such as intentional binding. Moreover, the putative volitional signal is not highly specific with respect to *which* agency relations are established. In our case, volitional signals controlling the right hand lead to intentional binding for involuntary movements of the left hand. Thus, a range of movement/outcome pairings may be enabled by co-occurrence with intention. Intentions do structure subsequent subjective experience, but by means of a loose fit, rather than a tight prediction about specific muscular movements. Previous studies suggested that the sense of agency is highly temporally-specific, in that intentions, actions, and outcomes must follow a predictable temporal sequence (Haggard and Clark, 2003; Bays et al., 2005). However, the *content* of intention, action and outcome can be combined arbitrarily without compromising the experience of agency. Our result suggests that volitional signals have the interesting property of high “latent associability”: they potentiate the development of *any* operant relation they co-occur with. This is consistent with Skinner’s demonstration that animals assume a causal connection between an action and a reinforcing stimulus, even when the connection is in fact an accidental correlation (Skinner, 1992).

In our case, the path between volition and agency is not effector-specific, but effector-independent. In particular, our design involved voluntary actions and involuntary movements assigned to different hands. Our results thus suggest that the contribution of internal volitional signals to sense of agency is bihemispheric, rather than hemisphere-specific. Rodent studies showed that mice readily learn to control a robot when arbitrary motor cortex activity is used to drive the robot dynamics. Learning such intentional neuroprosthetic skills depends on corticostriatal plasticity (Koralek et al., 2012). Our results likewise show that linking formation of an intention to an outcome leads to formation of some key temporal features of agency, even when the means that mediate between intention and outcome are artificial, and even after the original volitional signal is dropped. These findings may explain how humans can develop an enduring and successful feeling of agency in cases of non-natural movement like brain-machine interfaces (Nicolelis, 2003).

6.4.4. Learning one's own agency

Our experiment suggests that a conjunction of three conditions may be sufficient for sense of agency. First, an internal volitional signal must be present to provide a general metacognitive experience of intentional action. Second, some body movement must occur. Third, some external outcome of the action must occur. We also showed that no specific linkage between the metacognitive volitional signal and the body movement is necessary. In particular, the volitional signal need not be present at the same time as the body movement, nor even relate to the same effector. Thus, the internal volitional signal need not have a hardwired connection to the motor output system in the manner suggested by Sperry. In our experiment, it was sufficient that the volitional signal and the body movement had previously been associated.

This pattern of results reflects two fundamental features of human voluntary action, which we call *automaticity* and *flexibility*. *Automaticity* refers to the way that actions which initially require focussed attention, such as driving a car, or cooking soufflés, become increasingly

fluent with repetition. The subjective experience of action also changes. The action becomes less central in conscious experience, and instead provides a background 'buzz' of awareness (Synofzik et al., 2008). However outcomes are still fully attributed to one's own agency. Our results show a similar retention of key temporal features of experience of agency even when our experimental design deliberately reduced and removed intentional control over the outcome. Thus, our study can clarify a striking paradox of human action: namely, that one can feel fully in control of a skilled action such as riding a bicycle, and have a clear sense of agency, yet have only thin conscious experience of the action itself.

Flexibility refers to the ability of humans and animals to achieve control over goal states using complex and varying means (Engbert et al., 2008). This perhaps contributes to the astonishing human proficiency in developing and using technology. Hebb's classical concept of motor equivalence (Hebb, 2002) suggests that cognitive systems are not generally concerned whether a goal is achieved with one effector or with another – all movements that achieve the goal are effectively equivalent.

This transfer of key temporal features of experience of agency from intentional actions to other movements recalls the way that sense of agency emerges in human development. Human infants appear to act randomly, with little intentionality and goal-directedness, compared to healthy adults. During early experience, infants may gradually learn the precise mapping between different intentions, the resulting body movements, and external consequences. They thus eventually acquire the capacity to move a specific effector – achieving control over their body, and thus over their environment. Our results show that the capacity to form new intention-movement-outcome associations seems to remain and, importantly, could be generalised to non-voluntary movements, even when intentional action is no longer present. In this regard, it has been shown that younger children tend to confuse intended with accidental outcomes (Shultz et al., 1980; Shultz and Wells, 1985; Metcalfe et al., 2010).

Our experiments suggest that a hardwired internal volitional signal is required for the initial *acquisition* and emergence of sense agency. Importantly, this hardwired signal appears to be cognitive rather than motoric, since it is not linked to any specific output effector. At the same time, associative mechanisms contribute strongly to the *expression* of sense of agency. The presence of internal volitional signals during learning (path 1 in Fig.6.7.C) is necessary for induction, though not expression of key temporal features of experience for both direct voluntary action (path 2 in Fig.6.7.C), and also for an associated involuntary movements (path 3 in Fig.6.7.C).

Wittgenstein (Wittgenstein, 2009) famously asked “What is left over if I subtract the fact that my arm goes up from the fact that I raise my arm?”. Sense of agency is a partial answer to this question. However, even simple voluntary actions trigger widespread and automatic involuntary elements. For example, voluntarily lifting the right arm requires anticipatory compensations in contralateral muscles (Bouisset and Zattara, 1981; Schieber, 1995). Thus, voluntarily moving one effector normally leads to involuntary (or at least less voluntary) adjustments elsewhere, rather as in our experimental group. Importantly, people are not generally surprised, or even aware of these involuntary adjustments – although they would presumably be immediately conscious of a comparable passive displacement of the same body parts. Thus, the involuntary side-effects of voluntary action come to form part of an integrated experience of agency (Synofzik et al., 2008). The highly distributed, integrated nature of motor control ensures very frequent association between voluntary actions and involuntary movement. We suggest this fact lies at the heart of our finding of extensible sense of agency.

In conclusion, we suggest that some key temporal aspects of experience of agency, namely the perceptual anticipation of an action outcome, can be transferred from voluntary actions to involuntary movements. Such transfer follows repeated co-occurrence of an internal volitional signal, with both an involuntary body movement, and a sensory outcome. Importantly, association with an internal volitional signal appears to be necessary to initially

establish key temporal features of agency with respect to an involuntary movement, but is not necessary for its subsequent expression. The transfer process thus resembles the development of an enduring sense of agency that emerges during skill learning, as action control progresses from focussed and effortful to automatic. Interestingly, the involuntary movement that becomes associated need not match the intention precisely, suggesting that the metacognitive signals supporting agency acquisition are relatively non-specific. The high latent associability of these signals may reflect the distributed nature of motor control. Recent successes in acquisition of voluntary control using neuroprosthetics and brain-machine interfaces testify to the latent associability of human sense of agency.

Chapter 7

General Discussion

This thesis has explored the neural mechanisms underlying voluntary action control in the human brain and the experiences associated with it such as sense of agency. It has showed that voluntary actions arise from distinct cognitive processes (chapter 2) and that sense of agency could be a readout of these processes (chapter 3). It has provided new causal evidence that modulating specific parts of the brain that may host these processes such as parietal (chapter 4) and frontal (chapter 5) areas can influence the experiences associated with voluntary action. Finally, it has showed that by pairing voluntary actions of one hand with involuntary movements of the other hand, key aspects of agency experience can transfer from voluntary to involuntary movements, with repeated association (chapter 6). The final chapter aims to bring together the evidence presented in the previous chapters.

7.1. Is voluntary action noise?

New evidence suggest that the decision time to move in a self-initiated action is determined by the ebb and flow of background neuronal noise in the motor system, rather than the outcome of a specific neural event, as suggested by the classical RPs (Schurger et al., 2012; Murakami et al., 2014). We have showed in chapter 2 that in addition to stochastic fluctuations a gradual linear build-up of a preparatory process may also contribute to decision time to move in humans. However, the preparation itself may be triggered by a random event. Further, we have suggested that these precursor processes may be necessary for initiation of a self-initiated action, but may not be sufficient. Therefore, our study cannot show whether voluntary actions are caused by prior decisions, or by randomness, and our method cannot readily estimate the contribution of stochastic fluctuations to action decisions. Future experiments may be able to directly compare the magnitudes of stochastic fluctuations and fixed precursor processes to voluntary action generation. For example, by capturing the very low EEG frequency drifts that fail to trigger an action, non-phase-locked analyses such as background EEG spectral power could be a more powerful method to show the difference between self-initiated and externally-triggered actions compared to action-locked methods such as RPs or EEG SD decreases.

Importantly, the nature of this preparatory process remains unclear. As mentioned previously, the random processes in the brain may in fact be specific signals, if we could discover the entire range of complex interactions that govern the system. Therefore it would be a great challenge to dissociate the contribution of signal and noise in the brain to the timing of self-initiated actions.

One possibility is that a specific 'internal volitional signal' is fed into the motor system from higher-level neuronal circuits, which can switch the system into the 'volitional' mode to cause a self-initiated action and experiences associated with it such as agency and intention. Interestingly, Murakami et al. (2014) identified a subpopulation of neurons whose activity resembled to what one would expect from these 'internal volitional signals'. They were higher-level input to a stochastic accumulator with an inter-trial variability rate that was correlated with the amount of time the rat waited before performing a spontaneous action. This is not to claim that these signals are identical in both humans and rats, as we have no evidence yet that rodents can experience agency or intention. But our finding of a specific precursor process that kicks in about 1.5 s before action initiation strongly accords with a higher order input into an integrator, which can determine the exact time of action initiation.

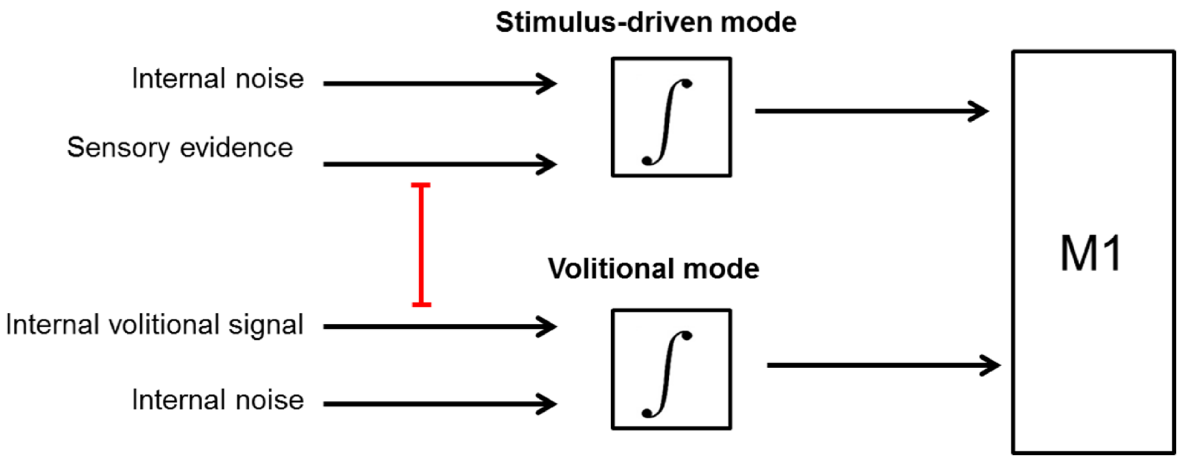


Fig.7.1. Specific 'internal volitional signals' may switch the action preparation system into 'volitional' mode, contributing to action initiation and subjective experiences associated with voluntary action. Our results from chapter 3 suggest that there might be a uni- or bidirectional inhibitory link between the internal and external mechanisms (red bold line).

On another account this preparatory process may be part of the motor noise itself which becomes part of the signal when an optimal subspace is reached. Based on this hypothesis, action initiation is preceded by a preparatory process during which stochastic fluctuations converge on a specific pattern. The drop in deviation from average activity will then be necessary to trigger execution of a voluntary action. However, this reduction in inter-trial variability could itself be triggered by random fluctuations or higher-order deterministic processes such as motivation (Manohar et al., 2015) or urgency (Cisek et al., 2009; Thura and Cisek, 2016).

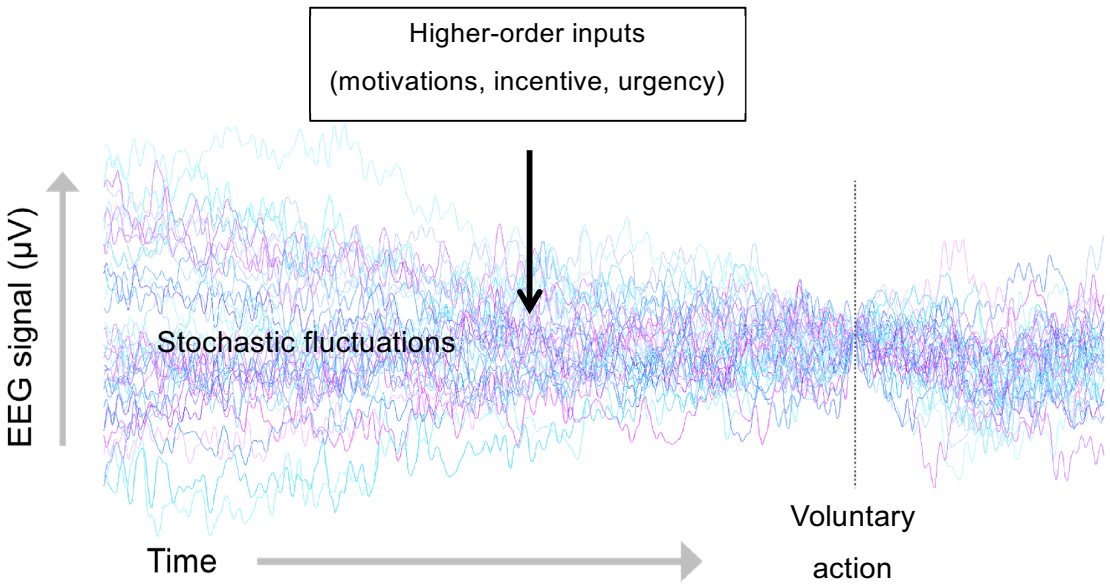


Fig.7.2. voluntary actions are preceded by a specific preparatory process in which inter-trial variability starts to converge. This event could itself be triggered by a random event or higher-level inputs such as motivation or urge.

Chapter 2 has showed that specific cognitive processes contribute to initiation of voluntary actions in addition to random fluctuations. However, self-initiated actions are distinguished from externally-triggered actions not only because they are internally generated but they are

also associated with specific subjective experiences such as experience of agency. Therefore, our findings from the second chapter bring on the possibility that the experiences of voluntary action could be a readout of these precursor cognitive processes.

7.2. Experience of voluntary action as a readout of internal precursor processes

Following our findings in chapter 2 we asked whether experience of agency could depend on a readout of brain processes in frontal areas that precede a voluntary action. If experience of agency depends on internal precursor signals, any intervention that influences these signals may also affect experience of agency. Chapter 3 has investigated this prediction by using a subliminal electrocutaneous stimulus as a probe. Moreover, we tested the same paradigm in a patient with anarchic hand syndrome (AHS) to better understand how the damage underlying AHS influences readout from these processes.

In accordance with our prediction, healthy participants experienced greater sense of agency when they were subliminally primed during action preparation. This suggested that sense of agency could not be purely retrospective. Rather sense of agency must depend, at least in part, on precursor cognitive processes during action preparation. Interestingly, this effect was not observed in a patient with AHS, suggesting that a mechanism that uses precursor signals of voluntary action to compute sense of agency was now disrupted.

Based on our findings in chapters 2 and 3 we have proposed a cognitive model of the experience of voluntary action (Fig.3.5). Based on this model one key input to the experience of agency is readout of internal volitional signals that precedes endogenous action. We have showed that this precursor process not only facilitates self-initiated action (chapter 2) but also contributes to the experiences associated with it (chapter 3). Further, in chapter 2 we proposed that the precursor process may be influenced by top down *internal*

signals such as motivations or desires. Chapter 3 has added to this hypothesis by showing that *external* environmental factors could also influence the processes preceding voluntary action. If the signals from the outside environment match with internal signals such as desire, they may integrate to facilitate action preparation and to construct a stronger experience of control over voluntary actions. Healthy participants can inhibit suggestions from the environment when they do not align with one's intention. However, chapter 3 has showed that this mechanism is impaired in patients with AHS. They frequently describe movements of the affected hand as involuntary, even when they are well-formed and co-ordinated, suggesting that patient's experience of actions is no longer driven by metacognitive readout of precursor processed preceding a voluntary action.

7.3. Modulating experiences of voluntary action with brain stimulation

If experiences of voluntary action are readout of brain processes in frontal and parietal areas, one should be able to influence these experiences by modulating the key brain circuits underlying the control of action. Building up on our previous findings, chapter 4 has investigated this hypothesis by modulating the activity in frontal areas by transcranial stimulation of DLPFC due to its role in action selection and preparation (Rowe et al., 2000). We also targeted angular gyrus (AG) based on previous studies that found activation in AG in conditions associated with reduced sense of agency (Farrer and Frith, 2002; Farrer et al., 2003a, 2008).

Participants experienced weaker sense of agency (as measured by intentional binding) when activity in AG was increased by anodal stimulation. We further showed that this effect is polarity and hemisphere specific. We hypothesised that the angular gyrus might contribute to the sense of agency by monitoring the linkage of actions to outcomes, or, alternatively and equivalently, failures of such linkage (Frith et al., 2000; Wolpert and Ghahramani, 2000). Anodal stimulation of this area may correspond to artificial boosting of a mismatch detection

process. We found no strong effect on experience of agency when modulating DLPFC. Previous chapters have showed that action initiation processes in frontal areas contribute to experiences of voluntary action. Therefore one may question why frontal anodal tDCS did not boost the sense of agency. Our task in chapter 4 did not manipulate action-outcome contingency and involved a simple habitual action to produce a single outcome. We therefore postulated that the contribution of precursor processes on experience of voluntary action may become more profound in a context where participants have to actively choose among several action alternatives.

Chapter 5 has investigated this possibility in a series of experiments manipulating a range of task parameters related to action selection and action outcomes, while participants were exposed to tDCS stimulation of the left DLPFC. Anodal stimulation of DLPFC increased our proxy measure of experience of agency, but only in tasks where participants endogenously selected between multiple actions. This result adds to our previous findings by showing that the contribution of action preparation processes in frontal areas to experiences of voluntary action is stronger when selecting between multiple actions. Finally, in a novel approach, chapter 5 has demonstrated applicability of meta-analysis of neurostimulation data to localise cognitive functions in the human brain.

7.4. Extending experiences of voluntary action

Having explored the neural correlates of voluntary action control and the experiences associated with it, chapter 6 has investigated the origin of this ubiquitous experience by asking whether sense of agency is hardwired, arising from precursor processes preceding a voluntary action within the brain's motor system (Fried et al., 2011), or rather depends on associative learning, through repeated association between a goal-directed action and another event (Prinz, 1997).

We showed that by pairing voluntary actions of one hand with involuntary movement of the other hand, key aspects of agency experience can transfer from voluntary to involuntary movements after repeated association. This finding suggest that an internal volitional signal is required for the primary acquisition of agency but, with repeated association, the involuntary movement in itself comes to produce some key temporal features of agency over the subsequent outcome. At first this seems in contrast with our previous claims that experiences of voluntary action is a readout of specific action preparatory processes in the brain that necessarily precedes a voluntary action. However, chapter 6 has showed that voluntary actions are necessary for the *initial* acquisition and emergence of a sense of agency. But once a voluntary signal is present, it can be mentally associated with other events, and spread to produce the distinctive intentional binding feature of volition, but now with respect to other movements.

Our findings in chapter 6 may also explain why as actions become less 'volitional' and more 'habitual', such as driving a car, their subjective experiences also change. Even though those actions are self-initiated and the outcomes are still fully attributed to one's own agency, they are more automatic, need less deliberation and are less responsive to reasons. This strategy, ability to perform voluntary actions without focal conscious attention, may reduce the computational cost in the motor system while retaining the main qualities of a voluntary action. Further, it demonstrates the underlying mechanism of another hallmark of volition, which is our ability to learn new skills and achieve control over goal states using complex and varying means.

Chapter 6 has complemented our previous findings from chapters 2 and 3 by showing that the metacognitive readout of precursor processes preceding a voluntary action is relatively nonspecific in regard to the outcome. This in turn suggests that the relation between precursor processes in the motor system and experiences associated with it is not precisely matched, and is not effector specific. Finally, by showing the transferability of experiences of

voluntary action from voluntary to involuntary movements, chapter 6 could guide acquisition of voluntary control using neuroprosthetics and brain–machine interfaces.

7.5. Future directions

Each of the chapters in this thesis has opened new possibilities for further investigations. For example, Chapter 2 has showed that stochastic fluctuations in the motor system are complemented by a specific precursor process to determine the exact time of a self-initiated action. However, it yet remains unclear to what extent does the endogenous decision of 'when' to move result from a specific neural precursor event, and to what extent does it result merely from background neural noise. Additionally, one may explore how environmental and internal physiological factors, including motivation, effort and reward can influence such endogenous, 'free', action decisions. We have also showed *'when'* in the brain these precursor processes start to emerge but not *'where'*. A further possibility is to investigate which brain areas house the first precursors of volitional action, and whether subjective experience of will could be caused by patterns of brain activity in these areas by using multivoxel pattern analysis of fMRI data. Finally, chapter 2 has concentrated on time domain analyses. However, RP could be viewed as a very-low frequency oscillation, although it is classically treated as single evoked potential. Investigation on the oscillatory effects of skip actions will further clarify the contribution of stochastic fluctuations to voluntary actions.

In chapter 6 we have showed that the experience of agency can *transfer* from voluntary to involuntary movements by repeated association. However, the mechanisms of this transfer remain unclear. An interesting avenue for further research is to use cortical surface reconstruction of individual structural MRI and EEG data to perform source reconstruction and connectivity analysis, in order to investigate whether interhemispheric functional connectivity plays a role in this transfer of agency experience. One might speculate that a

strong interhemispheric connectivity during the preparatory processes that precedes a voluntary action can lead to an increased experience of agency towards an outcome in subsequent involuntary movements.

7.6. Final conclusion

This thesis has provided a coherent groundwork on neural mechanisms underlying voluntary action control in the human brain. It has portrayed strong evidence in favour of a distinct system in the brain that underlies voluntary, self-initiated actions and the subjective experiences associated with it such as experience of agency. Importantly, these experiences are not hardwired, but depend on associative learning.

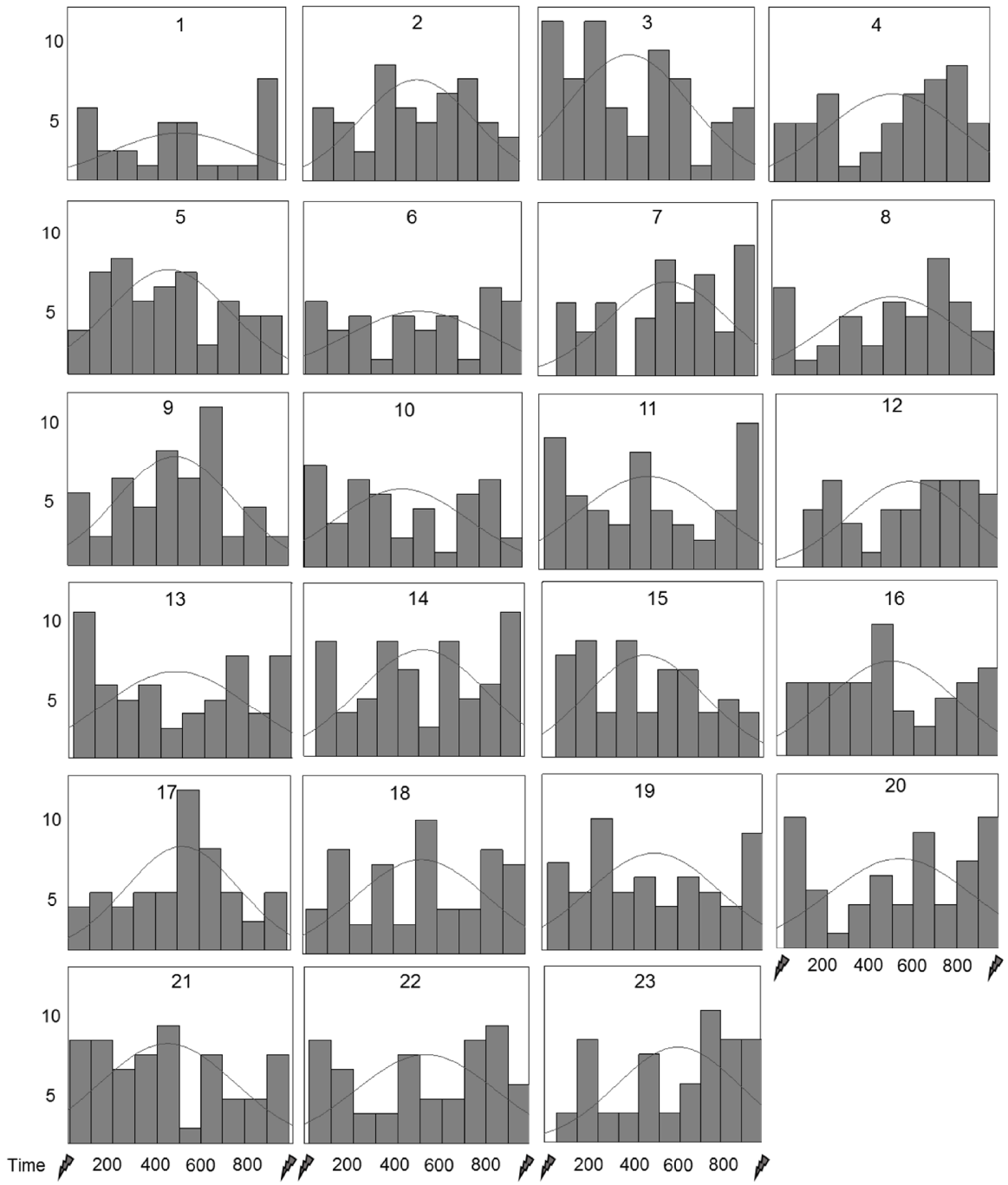
Our main findings could not be explained only based on reconstructive inference views on the experiences of will and agency (Wegner, 2003a) because preparatory brain events necessarily precedes both action and its effect. This thesis also complements new models on spontaneous, self-initiated actions (Schurger et al., 2012; Murakami et al., 2014), by showing that in addition to stochastic fluctuations a gradual linear build-up of a preparatory process may also contribute to decision time to move in humans. Finally, it has provided various evidence that experiences of voluntary action might be a metacognitive readout of the preparatory processes in frontal and parietal areas. Overall, we believe our findings rehabilitate volition as a cognitive concept.

Despite the undisputable importance of voluntary actions in social, moral and legal concepts, modern neuroscience has provided little actual evidence about the specific processes in the brain that make an action voluntary. This thesis has tried to identify neural basis of voluntary action by using rigorous experimental designs and analysis of brain data. The thesis outcomes could potentially have major implications for understanding the disorders of volition. They could provide a better mechanistic window into the origins of voluntary action, and could be a minor step towards a scientific understanding of what it means to be human.

Supplementary Data.

Subject	mean self-initiated	mean externally-triggered	SD self-initiated	SD externally-triggered
1	11.76	12.09	6.82	6.79
2	5.38	5.76	2.27	2.30
3	6.14	6.45	4.58	4.61
4	7.16	7.45	3.25	3.22
5	8.25	8.39	4.57	3.97
6	7.21	7.70	3.22	3.27
7	5.92	6.20	1.66	1.67
8	7.83	8.16	3.20	3.23
9	6.08	6.48	1.26	1.20
10	6.32	6.54	2.96	2.68
11	5.34	5.85	1.00	1.01
12	6.89	7.28	2.72	2.70
13	4.83	5.23	1.83	1.82
14	8.06	8.23	2.66	2.54
15	8.97	9.44	3.60	3.60
16	9.25	9.66	5.85	5.91
17	6.30	6.81	2.69	2.61
18	7.72	8.14	2.53	2.57
19	6.90	7.16	2.94	2.93
20	7.32	7.80	2.40	2.62
21	9.44	9.85	2.93	2.94
22	7.14	7.73	4.74	5.19

Supplementary table 2.1. Mean and distribution of waiting time before skipping in self-initiated and externally-triggered conditions. All values are in second.



Supplementary figure 3.1. Time histogram of latency of actions from their preceding subliminal shock on *primed-voluntary* trials. Participant's number is shown inside each plot.

Participant	Threshold (mA)	Subliminal shock intensity (mA)	Supraliminal shock intensity (mA)	Subliminal shock *d'*
1	0.43	0.42	0.55	0.92
2	0.37	0.36	0.47	1.44
3	0.46	0.45	0.59	1.12
4	0.54	0.53	0.69	0.68
5	0.26	0.25	0.33	0.68
6	0.43	0.42	0.55	1.12
7	0.41	0.40	0.52	1.29
8	0.47	0.46	0.60	1.12
9	0.33	0.32	0.42	1.12
10	0.70	0.69	0.90	0.61
11	0.53	0.52	0.68	0.68
12	0.44	0.43	0.56	1.29
13	0.70	0.69	0.90	0.68
14	0.35	0.34	0.44	1.12
15	0.37	0.36	0.47	1.44
16	0.48	0.47	0.61	0.92
17	0.55	0.54	0.70	0.92
18	0.57	0.56	0.73	1.12
19	0.41	0.40	0.52	1.12
20	0.72	0.71	0.92	0.97
21	0.48	0.47	0.61	1.12
22	1.13	1.12	1.46	1.12
23	0.44	0.43	0.56	1.12

Supplementary table 3.1. The shocks amplitudes and *d'* of each participant.

Participant	Primed-voluntary trials	Voluntary trials
1	16	-33
2	4	-31
3	39	28
4	-18	-45
5	33	-11
6	8	45
7	57	19
8	36	34
9	23	27
10	36	74
11	25	10
12	16	-8
13	13	-21
14	16	24
15	-22	-24
16	17	37
17	129	99
18	58	63
19	32	18
20	93	75
21	62	16
22	81	54
23	-24	-28

Supplementary table 3.2. Action binding (ms) in *primed-voluntary* and *voluntary* trials of healthy participants.

Participant	*p* value	Test statistic
1	0.81	0.44
2	0.58	0.68
3	0.03	2.99
4	0.23	1.32
5	0.37	0.98
6	0.39	0.94
7	0.04	2.60
8	0.39	0.94
9	0.00	Inf
10	0.30	1.11
11	0.63	0.61
12	0.03	2.92
13	0.33	1.05
14	0.58	0.68
15	0.43	0.88
16	0.00	Inf
17	0.26	1.23
18	0.76	0.48
19	0.00	Inf
20	0.14	1.67
21	0.18	1.50
22	0.55	0.71
23	0.01	3.98

Supplementary table 3.3. *p* value and test statistic of Anderson Darling test of the healthy participants. Shaded rows are participants with non-uniform keypress latency on *primed-voluntary* trials.

Trial number	Right hand Primed-voluntary trials	Right hand Voluntary trials	Left hand Primed-voluntary trials	Left hand Voluntary trials
1	124	37	3	170
2	284	19	124	-80
3	124	137	76	39
4	64	229	70	-117
5	305	250	-49	51
6	132	271	126	158
7	113	-24	44	116
8	27	203	89	-87
9	-45	64	1	269
10	-29	174	31	-21
11	124	55	-274	-18
12	63	-28	86	131
13	-53	162	38	44
14	147	50	-443	276
15	-142	-3	67	116
16	-19	-72	-437	48
17	45	170	0	-10
18	301	130	17	167
19	14	-33	118	62
20	97	159	-595	56
21	25	108	-90	-188
22	81	-37	-101	178
23	141	-61	-25	95
24	-90	-37	45	-65
25	396	-128		-61
26	401	219		-111
27	-38	-22		-118
28	20	-196		15
29	148	-103		13
30	147			95
31	92			-22
32	271			-66
33	31			
34	94			
35	67			
36	40			
37	195			

Supplementary table 3.4. Action binding (ms) in *primed-voluntary* and *voluntary* trials of the healthy (right) and the affected (left) hand of a patient with AHS.

References

Adams RA, Shipp S, Friston KJ (2013a) Predictions not commands: active inference in the motor system. Brain Struct Funct 218:611–643.

Adams RA, Stephan KE, Brown HR, Frith CD, Friston KJ (2013b) The computational anatomy of psychosis. Front Psychiatry 4:47.

Alexander P, Schlegel A, Sinnott-Armstrong W, Roskies AL, Wheatley T, Tse PU (2015) Readiness potentials driven by non-motoric processes. Conscious Cogn 39:38–47.

Anscombe GEM (2000) Intention. Harvard University Press.

Assal F, Schwartz S, Vuilleumier P (2007) Moving with or without will: functional neural correlates of alien hand syndrome. Ann Neurol 62:301–306.

Averbeck BB, Lee D (2003) Neural noise and movement-related codes in the macaque supplementary motor area. J Neurosci Off J Soc Neurosci 23:7630–7641.

Ball T, Schreiber A, Feige B, Wagner M, Lücking CH, Kristeva-Feige R (1999) The role of higher-order motor areas in voluntary movement as revealed by high-resolution EEG and fMRI. NeuroImage 10:682–694.

Bardi L, Kanai R, Mapelli D, Walsh V (2013) Direct current stimulation (tDCS) reveals parietal asymmetry in local/global and salience-based selection. Cortex 49:850–860.

Barlas Z, Obhi SS (2013) Freedom, choice, and the sense of agency. Front Hum Neurosci 7:514.

Bays PM, Wolpert DM, Flanagan JR (2005) Perception of the consequences of self-action is temporally tuned and event driven. Curr Biol CB 15:1125–1128.

Berlucchi G (2012) Frontal callosal disconnection syndromes. Cortex 48:36–45.

Blakemore SJ, Wolpert DM, Frith CD (2002) Abnormalities in the awareness of action. Trends Cogn Sci 6:237–242.

Blankenburg F, Taskin B, Ruben J, Moosmann M, Ritter P, Curio G, Villringer A (2003) Imperceptible stimuli and sensory processing impediment. Science 299:1864.

Bode S, Murawski C, Soon CS, Bode P, Stahl J, Smith PL (2014) Demystifying “free will”: The role of contextual information and evidence accumulation for predictive brain activity. Neurosci Biobehav Rev 47C:636–645.

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR (2011) Introduction to Meta-Analysis. John Wiley & Sons.

Boschert J, Hink RF, Deecke L (1983) Finger movement versus toe movement-related potentials: further evidence for supplementary motor area (SMA) participation prior to voluntary action. Exp Brain Res 52:73–80.

Bouisset S, Zattara M (1981) A sequence of postural movements precedes voluntary movement. Neurosci Lett 22:263–270.

Brainard DH (1997) The Psychophysics Toolbox. Spat Vis 10:433–436.

Brass M, Haggard P (2007) To do or not to do: the neural signature of self-control. J Neurosci Off J Soc Neurosci 27:9141–9145.

Browne M (2000) Cross-Validation Methods. J Math Psychol 44:108–132.

Brunia CHM, Damen EJP (1988) Distribution of slow brain potentials related to motor preparation and stimulus anticipation in a time estimation task. Electroencephalogr Clin Neurophysiol 69:234–243.

Buschke H, Fuld PA (1974) Evaluating storage, retention, and retrieval in disordered memory and learning. Neurology 24:1019–1025.

Buxbaum LJ, Sirigu A, Schwartz MF, Klatzky R (2003) Cognitive representations of hand posture in ideomotor apraxia. Neuropsychologia 41:1091–1113.

Bynum WF (1980) Sherrington. His life and thought. Med Hist 24:217.

Caffarra P, Vezzadini G, Dieci F, Zonato F, Venneri A (2014) Rey-Osterrieth complex figure: normative values in an Italian population sample. Neurol Sci 22:443–447.

Cardinal RN, Aitken MRF (2013) ANOVA for the Behavioral Sciences Researcher. Psychology Press.

Casini L, Vidal F (2011) The SMAs: neural substrate of the temporal accumulator? Front Integr Neurosci:35.

Caspar EA, Christensen JF, Cleeremans A, Haggard P (2016) Coercion Changes the Sense of Agency in the Human Brain. Curr Biol CB 26:585–592.

Cavazzana A, Penolazzi B, Begliomini C, Bisiacchi PS (2015) Neural underpinnings of the "agent brain": new evidence from transcranial direct current stimulation. Eur J Neurosci.

Chambon V, Haggard P (2012) Sense of control depends on fluency of action selection, not motor performance. Cognition 125:441–451.

Chambon V, Haggard P (2013) Premotor or Ideomotor: How Does the Experience of Action Come About? In: Action Science (Prinz W, Beisert M, Herwig A, eds), pp 358–380. The MIT Press. Available at: http://mitpress.universitypressscholarship.com/view/10.7551/mitpress/9780262018555.001.0001/upso-9780262018555-chapter-14 [Accessed August 24, 2016].

Chambon V, Moore JW, Haggard P (2014a) TMS stimulation over the inferior parietal cortex disrupts prospective sense of agency. Brain Struct Funct.

Chambon V, Sidarus N, Haggard P (2014b) From action intentions to action effects: how does the sense of agency come about? Front Hum Neurosci 8:320.

Chambon V, Wenke D, Fleming SM, Prinz W, Haggard P (2013) An Online Neural Substrate for Sense of Agency. Cereb Cortex 23:1031–1037.

Churchland AK, Kiani R, Chaudhuri R, Wang X-J, Pouget A, Shadlen MN (2011) Variance as a signature of neural computations during decision making. Neuron 69:818–831.

Churchland MM et al. (2010) Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nat Neurosci 13:369–378.

Churchland MM, Abbott LF (2012) Two layers of neural variability. Nat Neurosci 15:1472–1474.

Churchland MM, Yu BM, Ryu SI, Santhanam G, Shenoy KV (2006) Neural Variability in Premotor Cortex Provides a Signature of Motor Preparation. J Neurosci 26:3697–3712.

Cisek P (2007) A parallel framework for interactive behavior. Prog Brain Res 165:475–492.

Cisek P, Puskas GA, El-Murr S (2009) Decisions in changing conditions: the urgency-gating model. J Neurosci Off J Soc Neurosci 29:11560–11571.

Cohen Kadosh R, Soskic S, Iuculano T, Kanai R, Walsh V (2010) Modulating neuronal activity produces specific and long-lasting changes in numerical competence. Curr Biol CB 20:2016–2020.

Constantino SM, Daw ND (2015) Learning the opportunity cost of time in a patch-foraging task. Cogn Affect Behav Neurosci 15:837–853.

Cravo AM, Claessens PME, Baldo MVC (2009) Voluntary action and causality in temporal binding. Exp Brain Res 199:95–99.

Crawford JR, Garthwaite PH, Porter S (2010) Point and interval estimates of effect sizes for the case-controls design in neuropsychology: rationale, methods, implementations, and proposed reporting standards. Cogn Neuropsychol 27:245–260.

Cumming G (2005) Understanding the Average Probability of Replication Comment on Killeen (2005). Psychol Sci 16:1002–1004.

Cunnington R, Windischberger C, Deecke L, Moser E (2002) The preparation and execution of self-initiated and externally-triggered movement: a study of event-related fMRI. NeuroImage 15:373–385.

Cunnington R, Windischberger C, Deecke L, Moser E (2003) The preparation and readiness for voluntary movement: a high-field event-related fMRI study of the Bereitschafts-BOLD response. NeuroImage 20:404–412.

Cunnington R, Windischberger C, Moser E (2005) Premovement activity of the pre-supplementary motor area and the readiness for action: studies of time-resolved event-related functional MRI. Hum Mov Sci 24:644–656.

Damen E j. p., Brunia C h. m. (1987) Changes in Heart Rate and Slow Brain Potentials Related to Motor Preparation and Stimulus Anticipation in a Time Estimation Task. Psychophysiology 24:700–713.

David N, Cohen MX, Newen A, Bewernick BH, Shah NJ, Fink GR, Vogeley K (2007) The extrastriate cortex distinguishes between the consequences of one's own and others' behavior. NeuroImage 36:1004–1014.

De Renzi E, Motti F, Nichelli P (1980) Imitating gestures. A quantitative approach to ideomotor apraxia. Arch Neurol 37:6–10.

De Renzi E, Pieczuro A, Vignolo L (1968) Ideational Apraxia: A Quantitative Study. Neuropsychologia 6:41–52.

Deecke L, Kornhuber HH (1978) An electrical sign of participation of the mesial "supplementary" motor cortex in human voluntary finger movement. Brain Res 159:473–476.

Deiber MP, Honda M, Ibañez V, Sadato N, Hallett M (1999) Mesial motor areas in self-initiated versus externally triggered movements examined with fMRI: effect of movement type and rate. J Neurophysiol 81:3065–3077.

Della Sala S, Marchetti C, Spinnler H (1991) Right-sided anarchic (alien) hand: a longitudinal study. Neuropsychologia 29:1113–1127.

Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9–21.

Dennett D, Kinsbourne M (1992) Time and the Observer - the Where and When of Consciousness in the Brain. Behav Brain Sci 15:183–201.

Dennett DC (1985) Elbow Room: The Varieties of Free Will Worth Wanting. Oxford : New York: Oxford.

Desantis A, Waszak F, Gorea A (2016a) Agency alters perceptual decisions about action-outcomes. Exp Brain Res.

Desantis A, Waszak F, Moutsopoulou K, Haggard P (2016b) How action structures time: About the perceived temporal order of action and predicted outcomes. Cognition 146:100–109.

Desmurget M, Reilly KT, Richard N, Szathmari A, Mottolese C, Sirigu A (2009) Movement intention after parietal cortex stimulation in humans. Science 324:811–813.

Dirnberger G, Lang W, Lindinger G (2008) A new method to determine temporal variability in the period of pre-movement electroencephalographic activity. Int J Psychophysiol Off J Int Organ Psychophysiol 70:165–170.

Douglas ZH, Maniscalco B, Hallett M, Wassermann EM, He BJ (2015) Modulating conscious movement intention by noninvasive brain stimulation and the underlying neural mechanisms. J Neurosci Off J Soc Neurosci 35:7239–7255.

Drongelen W van (2006) Signal Processing for Neuroscientists: An Introduction to the Analysis of Physiological Signals, 1 Har/Cdr edition. Amsterdam ; Burlington, MA: Academic Press.

Dum RP, Strick PL (2002) Motor areas in the frontal lobe of the primate. Physiol Behav 77:677–682.

Ebert JP, Wegner DM (2010) Time warp: Authorship shapes the perceived timing of actions and events. Conscious Cogn 19:481–489.

Edwards MJ, Adams RA, Brown H, Pareés I, Friston KJ (2012) A Bayesian account of "hysteria." Brain J Neurol 135:3495–3512.

Efron B, Stein C (1981) The Jackknife Estimate of Variance. Ann Stat 9:586–596.

Eimer M, Schlaghecken F (1998) Effects of masked stimuli on motor activation: behavioral and electrophysiological evidence. J Exp Psychol Hum Percept Perform 24:1737–1747.

Eimer M, Schlaghecken F (2002) Links between conscious awareness and response inhibition: evidence from masked priming. Psychon Bull Rev 9:514–520.

Engbert K, Wohlschläger A, Haggard P (2008) Who is causing what? The sense of agency is relational and efferent-triggered. Cognition 107:693–704.

Engbert K, Wohlschläger A, Thomas R, Haggard P (2007) Agency, subjective time, and other minds. J Exp Psychol Hum Percept Perform 33:1261–1268.

Facchini S, Muellbacher W, Battaglia F, Boroojerdi B, Hallett M (2002) Focal enhancement of motor cortex excitability during motor imagery: a transcranial magnetic stimulation study. Acta Neurol Scand 105:146–151.

Fadiga L, Buccino G, Craighero L, Fogassi L, Gallese V, Pavesi G (1999) Corticospinal excitability is specifically modulated by motor imagery: a magnetic stimulation study. Neuropsychologia 37:147–158.

Faisal AA, Selen LPJ, Wolpert DM (2008) Noise in the nervous system. Nat Rev Neurosci 9:292–303.

Farrer C, Franck N, Georgieff N, Frith CD, Decety J, Jeannerod M (2003a) Modulating the experience of agency: a positron emission tomography study. NeuroImage 18:324–333.

Farrer C, Franck N, Paillard J, Jeannerod M (2003b) The role of proprioception in action recognition. Conscious Cogn 12:609–619.

Farrer C, Frey SH, Horn JDV, Tunik E, Turk D, Inati S, Grafton ST (2008) The Angular Gyrus Computes Action Awareness Representations. Cereb Cortex 18:254–261.

Farrer C, Frith CD (2002) Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. NeuroImage 15:596–603.

Feinberg TE, Schindler RJ, Flanagan NG, Haber LD (1992) Two alien hand syndromes. Neurology 42:19–19.

Filevich E, Kühn S, Haggard P (2012) Negative motor phenomena in cortical stimulation: implications for inhibitory control of human action. Cortex J Devoted Study Nerv Syst Behav 48:1251–1261.

Filmer HL, Dux PE, Mattingley JB (2014) Applications of transcranial direct current stimulation for understanding brain function. Trends Neurosci.

Fink GR, Marshall JC, Halligan PW, Frith CD, Driver J, Frackowiak RS, Dolan RJ (1999) The neural consequences of conflict between intention and the senses. Brain J Neurol 122 (Pt 3):497–512.

Fitts PM (1954) The information capacity of the human motor system in controlling the amplitude of movement. J Exp Psychol 47:381–391.

Fleming SM, Frith CD (2014) The Cognitive Neuroscience of Metacognition. Springer Science & Business Media.

Fletcher PC, Frith CD (2009) Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. Nat Rev Neurosci 10:48–58.

Fried I, Katz A, McCarthy G, Sass KJ, Williamson P, Spencer SS, Spencer DD (1991) Functional organization of human supplementary motor cortex studied by electrical stimulation. J Neurosci 11:3656–3666.

Fried I, Mukamel R, Kreiman G (2011) Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. Neuron 69:548–562.

Friston K (2010) The free-energy principle: a unified brain theory? Nat Rev Neurosci 11:127–138.

Friston K (2011) What is optimal about motor control? Neuron 72:488–498.

Frith CD (2014) Action, agency and responsibility. Neuropsychologia 55:137–142.

Frith CD, Blakemore, Wolpert DM (2000) Abnormalities in the awareness and control of action. Philos Trans R Soc Lond B Biol Sci 355:1771–1788.

Fukushima H, Goto Y, Maeda T, Kato M, Umeda S (2013) Neural Substrates for Judgment of Self-Agency in Ambiguous Situations. PLOS ONE 8:e72267.

Ganos C, Asmuss L, Bongert J, Brandt V, Münchau A, Haggard P (2015) Volitional action as perceptual detection: predictors of conscious intention in adolescents with tic disorders. Cortex J Devoted Study Nerv Syst Behav 64:47–54.

Gibson JJ (1986) The Ecological Approach to Visual Perception, New Ed edition. Boston: Routledge.

Gold JI, Shadlen MN (2007) The Neural Basis of Decision Making. Annu Rev Neurosci 30:535–574.

Gordon M, Lumley T (2015) forestplot: Advanced Forest Plot Using “grid” Graphics. R package version 1.3. Available at: http://CRAN.R-project.org/package=forestplot.

Haggard P (2005) Conscious intention and motor cognition. Trends Cogn Sci 9:290–295.

Haggard P (2008) Human volition: towards a neuroscience of will. Nat Rev Neurosci 9:934–946.

Haggard P (2011) Decision Time for Free Will. Neuron 69:404–406.

Haggard P, Aschersleben G, Gehrke J, Prinz W (2002a) Action, binding, and awareness. In: Common mechanisms in perception and action B (Prinz W, Hommel B, eds), pp

266–285. OXFORD UNIVERSITY PRESS. Available at: http://eprints.ucl.ac.uk/4192/ [Accessed August 27, 2014].

Haggard P, Chambon V (2012) Sense of agency. Curr Biol CB 22:R390-392.

Haggard P, Clark S (2003) Intentional action: Conscious experience and neural prediction. Conscious Cogn 12:695–707.

Haggard P, Clark S, Kalogeras J (2002b) Voluntary action and conscious awareness. Nat Neurosci 5:382–385.

Haggard P, Clark S, Kalogeras J (2002c) Voluntary action and conscious awareness. Nat Neurosci 5:382–385.

Haggard P, Eimer M (1999) On the relation between brain potentials and the awareness of voluntary movements. Exp Brain Res 126:128–133.

Haggard P in the I of CN and the D of PP, Eitam B (2015) The Sense of Agency. Oxford University Press, USA.

Haggard P, Martin F, Taylor-Clarke M, Jeannerod M, Franck N (2003) Awareness of action in schizophrenia. Neuroreport 14:1081–1085.

Hämmerer D, Bonaiuto J, Klein-Flügge M, Bikson M, Bestmann S (2016) Selective alteration of human value decisions with medial frontal tDCS is predicted by changes in attractor dynamics. Sci Rep 6:25160.

Hart HLA, Honore T (1985) Causation in the Law, Second edition. Oxford : New York: Oxford University Press, USA.

Hassan A, Josephs KA (2016) Alien Hand Syndrome. Curr Neurol Neurosci Rep 16:73.

Hauser M, Moore JW, de Millas W, Gallinat J, Heinz A, Haggard P, Voss M (2011) Sense of agency is altered in patients with a putative psychotic prodrome. Schizophr Res 126:20–27.

He BJ (2013) Spontaneous and task-evoked brain activity negatively interact. J Neurosci Off J Soc Neurosci 33:4672–4682.

Heaton R, Chelune G, Talley J, Kay G, Curtiss G (2000) WCST: Wisconsin Card Sorting test. Italian Edition edited by Hardoy M.C., Carta M.G., Hardoy M.J., Cabras P.L. Firenze: Giunti-Organizzazioni Speciali.

Hebb DO (2002) The Organization of Behavior: A Neuropsychological Theory, New Ed edition. Mahwah, N.J: Psychology Press.

Hecht D, Walsh V, Lavidor M (2010) Transcranial Direct Current Stimulation Facilitates Decision Making in a Probabilistic Guessing Task. J Neurosci 30:4241–4245.

Hedges LV, Vevea JL (1998) Fixed- and random-effects models in meta-analysis. Psychol Methods 3:486–504.

Higgins JPT, Green S (n.d.) 9.5.2 Identifying and measuring heterogeneity. Available at: http://handbook.cochrane.org/chapter_9/9_5_2_identifying_and_measuring_heterogeneity.htm [Accessed August 10, 2015a].

Higgins JPT, Green S (n.d.) 9.6.4 Meta-regression. Available at: http://handbook.cochrane.org/chapter_9/9_6_4_meta_regression.htm [Accessed January 2, 2016b].

Higgins JPT, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. BMJ 327:557–560.

Horvath JC, Forte JD, Carter O (2014) Evidence that transcranial direct current stimulation (tDCS) generates little-to-no reliable neurophysiologic effect beyond MEP amplitude modulation in healthy human subjects: A systematic review. Neuropsychologia 66C:213–236.

Horvath JC, Forte JD, Carter O (2015a) Transcranial direct current stimulation demonstrates little-to-no reliable or significant effect on any cognitive/behavioral or neurophysiologic outcome measure: a comprehensive meta-analysis. Brain Stimul Basic Transl Clin Res Neuromodulation 8:318–319.

Horvath JC, Forte JD, Carter O (2015b) Quantitative Review Finds No Evidence of Cognitive Effects in Healthy Populations From Single-session Transcranial Direct Current Stimulation (tDCS). Brain Stimulat 8:535–550.

Hughes G, Schütz-Bosbach S, Waszak F (2011) One Action System or Two? Evidence for Common Central Preparatory Mechanisms in Voluntary and Stimulus-Driven Actions. J Neurosci 31:16692–16699.

Jacobson L, Goren N, Lavidor M, Levy DA (2012a) Oppositional transcranial direct current stimulation (tDCS) of parietal substrates of attention during encoding modulates episodic memory. Brain Res 1439:66–72.

Jacobson L, Koslowsky M, Lavidor M (2012b) tDCS polarity effects in motor and cognitive domains: a meta-analytical review. Exp Brain Res 216:1–10.

Jahanshahi M (2013) The Bereitschaftspotential: Movement-Related Cortical Potentials, Softcover reprint of the original 1st ed. 2003 edition. Boston, MA: Springer.

Jahanshahi M, Dirnberger G (1999) The left dorsolateral prefrontal cortex and random generation of responses: studies with transcranial magnetic stimulation. Neuropsychologia 37:181–190.

Jahanshahi M, Jenkins IH, Brown RG, Marsden CD, Passingham RE, Brooks DJ (1995) Self-initiated versus externally triggered movements. I. An investigation using measurement of regional cerebral blood flow with PET and movement-related potentials in normal and Parkinson's disease subjects. Brain J Neurol 118 (Pt 4):913–933.

Jenkins IH, Jahanshahi M, Jueptner M, Passingham RE, Brooks DJ (2000) Self-initiated versus externally triggered movements. II. The effect of movement predictability on regional cerebral blood flow. Brain J Neurol 123 (Pt 6):1216–1228.

Jo H-G, Hinterberger T, Wittmann M, Borghardt TL, Schmidt S (2013) Spontaneous EEG fluctuations determine the readiness potential: is preconscious brain activation a preparation process to move? Exp Brain Res 231:495–500.

Kagaya K, Takahata M (2010) Readiness discharge for spontaneous initiation of walking in crayfish. J Neurosci Off J Soc Neurosci 30:1348–1362.

Khalighinejad N, Haggard P (2015) Modulating human sense of agency with non-invasive brain stimulation. Cortex J Devoted Study Nerv Syst Behav 69:93–103.

Kiesel A, Miller J, Jolicoeur P, Brisson B (2008) Measurement of ERP latency differences: a comparison of single-participant and jackknife-based scoring methods. Psychophysiology 45:250–274.

Kim IH, Lee S, Lee C-Y, Lee DG (2014) Intracranial Hemorrhage in the Corpus Callosum Presenting as Callosal Disconnection Syndrome: FDG-PET and Tractography: A Case Report. Ann Rehabil Med 38:871–875.

Kloesel B, Czarnecki K, Muir JJ, Keller AS (2010) Sequelae of a left-sided parietal stroke: posterior alien hand syndrome. Neurocase 16:488–493.

Kolling N, Behrens TEJ, Mars RB, Rushworth MFS (2012) Neural mechanisms of foraging. Science 336:95–98.

Koralek AC, Jin X, Long JD, Costa RM, Carmena JM (2012) Corticostriatal plasticity is necessary for learning intentional neuroprosthetic skills. Nature 483:331–335.

Kornhuber HH, Deecke L (1965) Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. Pflüg Arch Für Gesamte Physiol Menschen Tiere 284:1–17.

Kornhuber HH, Deecke L (1990) Readiness for movement - The Bereitschaftspotential story. Citation Classic Commentary. Curr Contents Life Sci 33.

Kranick SM, Hallett M (2013) Neurology of volition. Exp Brain Res 229:313–327.

Krieghoff V, Waszak F, Prinz W, Brass M (2011) Neural and behavioral correlates of intentional actions. Neuropsychologia 49:767–776.

Krzanowski WJ (2000) Principles of Multivariate Analysis: A User's Perspective, 2 edition. Oxford Oxfordshire ; New York: OUP Oxford.

Kühn S, Brass M, Haggard P (2013) Feeling in control: Neural correlates of experience of agency. Cortex 49:1935–1942.

Lakens D (2013) Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Cognition 4:863.

Lau HC, Rogers RD, Haggard P, Passingham RE (2004) Attention to intention. Science 303:1208–1210.

Lee W, Reeve J (2013) Self-determined, but not non-self-determined, motivation predicts activations in the anterior insular cortex: an fMRI study of personal agency. Soc Cogn Affect Neurosci 8:538–545.

Li S, Stevens JA, Rymer WZ (2009) Interactions between imagined movement and the initiation of voluntary movement: a TMS study. Clin Neurophysiol Off J Int Fed Clin Neurophysiol 120:1154–1160.

Libet B, Gleason CA, Wright EW, Pearl DK (1983) Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. Brain J Neurol 106 (Pt 3):623–642.

Lipsey MW, Wilson DB (2001) Practical Meta-Analysis. SAGE Publications.

Litwin-Kumar A, Doiron B (2012) Slow dynamics and high variability in balanced cortical networks with clustered connections. Nat Neurosci 15:1498–1505.

Llinas P (2002) I of the Vortex: From Neurons to Self, New Ed edition. MIT Press.

Luck SJ (2005) An Introduction to the Event-Related Potential Technique. Cambridge, Mass: MIT Press.

Manohar SG, Chong TT-J, Apps MAJ, Batla A, Stamelou M, Jarman PR, Bhatia KP, Husain M (2015) Reward Pays the Cost of Noise Reduction in Motor and Cognitive Control. Curr Biol CB 25:1707–1716.

Maoz U, Yaffe G (2014) Cognitive Neuroscience and Criminal Responsibility. In: The Cognitive Neurosciences, 5th Revised edition edition. Cambridge, Massachusetts: MIT Press.

Marcos E, Pani P, Brunamonti E, Deco G, Ferraina S, Verschure P (2013) Neural variability in premotor cortex is modulated by trial history and predicts behavioral performance. Neuron 78:249–255.

McBride J, Sumner P, Jackson SR, Bajaj N, Husain M (2013) Exaggerated object affordance and absent automatic inhibition in alien hand syndrome. Cortex J Devoted Study Nerv Syst Behav 49:2040–2054.

Meier U (2006) A note on the power of Fisher's least significant difference procedure. Pharm Stat 5:253–263.

Meiron O, Lavidor M (2013) Unilateral Prefrontal Direct Current Stimulation Effects are Modulated by Working Memory Load and Gender. Brain Stimulat 6:440–447.

Meron D, Hedger N, Garner M, Baldwin DS (n.d.) Transcranial direct current stimulation (tDCS) in the treatment of depression: systematic review and meta-analysis of efficacy and tolerability. Neurosci Biobehav Rev Available at: http://www.sciencedirect.com/science/article/pii/S0149763415001979 [Accessed August 10, 2015].

Metcalfe J, Eich TS, Castel AD (2010) Metacognition of agency across the lifespan. Cognition 116:267–282.

Metcalfe J, Greene MJ (2007) Metacognition of agency. J Exp Psychol Gen 136:184–199.

Moore JW, Fletcher PC (2012) Sense of agency in health and disease: a review of cue integration approaches. Conscious Cogn 21:59–68.

Moore JW, Lagnado D, Deal DC, Haggard P (2009a) Feelings of control: Contingency determines experience of action. Cognition 110:279–283.

Moore JW, Middleton D, Haggard P, Fletcher PC (2012) Exploring implicit and explicit aspects of sense of agency. Conscious Cogn 21:1748–1753.

Moore JW, Obhi SS (2012) Intentional binding and the sense of agency: a review. Conscious Cogn 21:546–561.

Moore JW, Ruge D, Wenke D, Rothwell J, Haggard P (2010) Disrupting the experience of control in the human brain: pre-supplementary motor area contributes to the sense of agency. Proc Biol Sci 277:2503–2509.

Moore JW, Wegner DM, Haggard P (2009b) Modulating the sense of agency with external cues. Conscious Cogn 18:1056–1064.

Moretto G, Walsh E, Haggard P (2011) Experience of agency and sense of responsibility. Conscious Cogn 20:1847–1854.

Morton J, Marcus S, Frankish C (1976) Perceptual centers (P-centers). Psychol Rev 83:405–408.

Moutard C, Dehaene S, Malach R (2015) Spontaneous Fluctuations and Non-linear Ignitions: Two Dynamic Faces of Cortical Recurrent Loops. Neuron 88:194–206.

Murakami M, Vicente MI, Costa GM, Mainen ZF (2014) Neural antecedents of self-initiated actions in secondary motor cortex. Nat Neurosci.

Nachev P, Husain M (2010) Action and the fallacy of the “internal”: Comment on Passingham et al. Trends Cogn Sci 14:192-193-194.

Nachev P, Kennard C, Husain M (2008) Functional role of the supplementary and pre-supplementary motor areas. Nat Rev Neurosci 9:856–869.

Nahab FB, Kundu P, Gallea C, Kakareka J, Pursley R, Pohida T, Miletta N, Friedman J, Hallett M (2011) The neural processes underlying self-agency. Cereb Cortex N Y N 1991 21:48–55.

Nathan SS, Sinha SR, Gordon B, Lesser RP, Thakor NV (1993) Determination of current density distributions generated by electrical stimulation of the human cerebral cortex. Electroencephalogr Clin Neurophysiol 86:183–192.

Nicolelis MAL (2003) Brain–machine interfaces to restore motor function and probe neural circuits. Nat Rev Neurosci 4:417–422.

Nitsche MA, Bikson M, Bestmann S (2015) On the Use of Meta-analysis in Neuromodulatory Non-invasive Brain Stimulation. Brain Stimulat 8:666–667.

Nitsche MA, Cohen LG, Wassermann EM, Priori A, Lang N, Antal A, Paulus W, Hummel F, Boggio PS, Fregni F, Pascual-Leone A (2008a) Transcranial direct current stimulation: State of the art 2008. Brain Stimulat 1:206–223.

Nitsche MA, Cohen LG, Wassermann EM, Priori A, Lang N, Antal A, Paulus W, Hummel F, Boggio PS, Fregni F, Pascual-Leone A (2008b) Transcranial direct current stimulation: State of the art 2008. Brain Stimulat 1:206–223.

Nitsche MA, Liebetanz D, Lang N, Antal A, Tergau F, Paulus W (2003) Safety criteria for transcranial direct current stimulation (tDCS) in humans. Clin Neurophysiol Off J Int Fed Clin Neurophysiol 114:2220-2222-2223.

Nitsche MA, Paulus W (2000) Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. J Physiol 527 Pt 3:633–639.

Obhi SS, Haggard P (2004) Internally generated and externally triggered actions are physically distinct and independently controlled. Exp Brain Res 156:518–523.

Obhi SS, Hall P (2011) Sense of agency and intentional binding in joint action. Exp Brain Res 211:655–662.

Obhi SS, Matkovich S, Chen R (2009) Changing the “When” and “What” of Intended Actions. J Neurophysiol 102:2755–2762.

Okamoto M, Dan H, Sakamoto K, Takeo K, Shimizu K, Kohno S, Oda I, Isobe S, Suzuki T, Kohyama K, Dan I (2004) Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10–20 system oriented for transcranial functional brain mapping. NeuroImage 21:99–111.

Okano K, Tanji J (1987) Neuronal activities in the primate motor fields of the agranular frontal cortex preceding visually triggered and self-paced movement. Exp Brain Res 66:155–166.

Oostenveld R, Fries P, Maris E, Schoffelen J-M, Oostenveld R, Fries P, Maris E, Schoffelen J-M (2010) FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data, FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. Comput Intell Neurosci Comput Intell Neurosci 2011, 2011:e156869.

Orsini A, Grossi D, Capitani E, Laiacona M, Papagno C, Vallar DG (1987) Verbal and spatial immediate memory span: Normative data from 1355 adults and 1112 children. Ital J Neurol Sci 8:537–548.

Osterrieth P (1944) Le test de copie d’une figure complexe; contribution à l’étude de la perception et de la mémoire. [Test of copying a complex figure; contribution to the study of perception and memory.]. Arch Psychol 30:206–356.

Papa SM, Artieda J, Obeso JA (1991) Cortical activity preceding self-initiated and externally triggered voluntary movement. Mov Disord Off J Mov Disord Soc 6:217–224.

Passingham R (1995) The Frontal Lobes and Voluntary Action. Oxford University Press.

Passingham RE (1987) Two cortical systems for directing movement. Ciba Found Symp 132:151–164.

Passingham RE, Bengtsson SL, Lau HC (2010a) Medial frontal cortex: from self-generated action to reflection on one’s own performance. Trends Cogn Sci 14:16–21.

Passingham RE, Bengtsson SL, Lau HC (2010b) Is it fallacious to talk of self-generated action?: Response to Nachev and Husain. Trends Cogn Sci 14:193–194.

Pavlov (1927) PI (2010) Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. Ann Neurosci 17:136–141.

Pereboom D (2011) Free Will Skepticism and Meaning in Life. In: The Oxford Handbook of Free Will (Kane R, ed). Oxford University Press.

Pink T (2004) Free Will: A Very Short Introduction. Oxford ; New York: OUP Oxford.

Poreisz C, Boros K, Antal A, Paulus W (2007) Safety aspects of transcranial direct current stimulation concerning healthy subjects and patients. Brain Res Bull 72:208–214.

Prinz W (1997) Perception and Action Planning. Eur J Cogn Psychol 9:129–154.

Pronin E, Wegner DM, McCarthy K, Rodriguez S (2006) Everyday magical powers: the role of apparent mental causation in the overestimation of personal influence. J Pers Soc Psychol 91:218–231.

Raja Beharelle A, Polanía R, Hare TA, Ruff CC (2015) Transcranial Stimulation over Frontopolar Cortex Elucidates the Choice Attributes and Neural Mechanisms Used to Resolve Exploration-Exploitation Trade-Offs. J Neurosci Off J Soc Neurosci 35:14544–14556.

Reale G (1985) The Systems of the Hellenistic Age: History of Ancient Philosophy (Suny Series in Philosophy): Systems of the Hellenistic Age Vol 3 (Catan JR, ed). Albany: State University of New York Press.

Reis J, Fritsch B (2011) Modulation of motor performance and motor learning by transcranial direct current stimulation. Curr Opin Neurol 24:590–596.

Riddoch MJ, Edwards MG, Humphreys GW, West R, Heafield T (1998) Visual affordances direct action: neuropsychological evidence from manual interference. Cogn Neuropsychol 15:645–683.

Rorden C, Brett M (2000) Stereotaxic display of brain lesions. Behav Neurol 12:191–200.

Rorden C, Karnath H-O, Bonilha L (2007) Improving lesion-symptom mapping. J Cogn Neurosci 19:1081–1088.

Rossi S, Hallett M, Rossini PM, Pascual-Leone A (2009) Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. Clin Neurophysiol 120:2008–2039.

Rowe JB, Hughes L, Nimmo-Smith I (2010) Action selection: a race model for selected and non-selected actions distinguishes the contribution of premotor and prefrontal areas. NeuroImage 51:888–896.

Rowe JB, Toni I, Josephs O, Frackowiak RS, Passingham RE (2000) The prefrontal cortex: response selection or maintenance within working memory? Science 288:1656–1660.

Ryle G, Dennett DC (2000) The Concept of Mind, 1 edition. Chicago: University Of Chicago Press.

Saberi-Moghadam S, Ferrari-Toniolo S, Ferraina S, Caminiti R, Battaglia-Mayer A (2016) Modulation of Neural Variability in Premotor, Motor, and Posterior Parietal Cortex during Change of Motor Intention. J Neurosci 36:4614–4623.

Saito N, Takahata K, Murai T, Takahashi H (2015) Discrepancy between explicit judgement of agency and implicit feeling of agency: Implications for sense of agency and its disorders. Conscious Cogn 37:1–7.

Schieber MH (1995) Muscular production of individuated finger movements: the roles of extrinsic finger muscles. J Neurosci Off J Soc Neurosci 15:284–297.

Schlegel A, Alexander P, Sinnott-Armstrong W, Roskies A, Tse PU, Wheatley T (2013) Barking up the wrong free: readiness potentials reflect processes independent of conscious will. Exp Brain Res 229:329–335.

Schmidt S, Jo H-G, Wittmann M, Hinterberger T (2016) “Catching the Waves” - Slow Cortical Potentials as Moderator of Voluntary Action. Neurosci Biobehav Rev.

Schultze-Kraft M, Birman D, Rusconi M, Allefeld C, Görgen K, Dähne S, Blankertz B, Haynes J-D (2015) The point of no return in vetoing self-initiated movements. Proc Natl Acad Sci U S A.

Schurger A, Sarigiannidis I, Naccache L, Sitt JD, Dehaene S (2015) Cortical activity is more stable when sensory stimuli are consciously perceived. Proc Natl Acad Sci U S A 112:E2083-2092.

Schurger A, Sitt JD, Dehaene S (2012) An accumulator model for spontaneous neural activity prior to self-initiated movement. Proc Natl Acad Sci 109:E2904–E2913.

Schüür F, Haggard P (2011) What are self-generated actions? Conscious Cogn 20:1697–1704.

Shadlen MN, Gold JI (2004) In: The Cognitive Neurosciences, 3rd edn. (Gazzaniga MS, ed), pp 1229–1241. MIT Press.

Shadlen MN, Newsome WT (1998) The Variable Discharge of Cortical Neurons: Implications for Connectivity, Computation, and Information Coding. J Neurosci 18:3870–3896.

Shallice T (1988) From Neuropsychology to Mental Structure. Cambridge, UK: Cambridge University Press.

Shenoy KV, Kaufman MT, Sahani M, Churchland MM (2011) A dynamical systems view of motor preparation: implications for neural prosthetic system design. Prog Brain Res 192:33–58.

Shergill SS, Bays PM, Frith CD, Wolpert DM (2003) Two eyes for an eye: the neuroscience of force escalation. Science 301:187.

Sherrington CS (1906) The integrative action of the nervous system. New York, C Scribner's sons. Available at: http://archive.org/details/integrativeacti02shergoog [Accessed October 17, 2015].

Shibasaki H, Barrett G, Halliday E, Halliday AM (1980) Components of the movement-related cortical potential and their scalp topography. Electroencephalogr Clin Neurophysiol 49:213–226.

Shibasaki H, Hallett M (2006) What is the Bereitschaftspotential? Clin Neurophysiol Off J Int Fed Clin Neurophysiol 117:2341–2356.

Shu Y, Hasenstaub A, Badoual M, Bal T, McCormick DA (2003) Barrages of synaptic activity control the gain and sensitivity of cortical neurons. J Neurosci Off J Soc Neurosci 23:10388–10401.

Shultz TR, Wells D (1985) Judging the intentionality of action-outcomes. Dev Psychol 21:83–89.

Shultz TR, Wells D, Sarda M (1980) Development of the ability to distinguish intended actions from mistakes, reflexes, and passive movements. Br J Soc Clin Psychol 19:301–310.

Singer T, Seymour B, O'Doherty JP, Stephan KE, Dolan RJ, Frith CD (2006) Empathic neural responses are modulated by the perceived fairness of others. Nature 439:466–469.

Sirigu A, Daprati E, Ciancia S, Giraux P, Nighoghossian N, Posada A, Haggard P (2004) Altered awareness of voluntary action after damage to the parietal cortex. Nat Neurosci 7:80–84.

Sirigu A, Daprati E, Pradat-Diehl P, Franck N, Jeannerod M (1999) Perception of self-generated movement following left parietal lesion. Brain 122:1867–1874.

Skinner BF (1992) "Superstition" in the pigeon. 1948. J Exp Psychol Gen 121:273–274.

Soon CS, Brass M, Heinze H-J, Haynes J-D (2008) Unconscious determinants of free decisions in the human brain. Nat Neurosci 11:543–545.

Sperduti M, Delaveau P, Fossati P, Nadel J (2011) Different brain structures related to self- and external-agency attribution: a brief review and meta-analysis. Brain Struct Funct 216:151–157.

Sperry RW (1940) The functional results of muscle transposition in the hind limb of the rat. J Comp Neurol 73:379–404.

Sperry RW (1943) Effect of 180 degree rotation of the retinal field on visuomotor coordination. J Exp Zool 92:263–279.

Spinnler H, Tognoni G (1987) [Italian standardization and classification of Neuropsychological tests. The Italian Group on the Neuropsychological Study of Aging]. Ital J Neurol Sci Suppl 8:1–120.

Spitoni GF, Pireddu G, Cimmino RL, Galati G, Priori A, Lavidor M, Jacobson L, Pizzamiglio L (2013) Right but not left angular gyrus modulates the metric component of the mental body representation: a tDCS study. Exp Brain Res Exp Hirnforsch Expérimentation Cérébrale 228:63–72.

Stagg CJ, Nitsche MA (2011) Physiological basis of transcranial direct current stimulation. Neurosci Rev J Bringing Neurobiol Neurol Psychiatry 17:37–53.

Stein RB, Gossen ER, Jones KE (2005) Neuronal variability: noise or part of the signal? Nat Rev Neurosci 6:389–397.

Sumner P, Husain M (2008) At the edge of consciousness: automatic motor activation and voluntary control. Neurosci Rev J Bringing Neurobiol Neurol Psychiatry 14:474–486.

Sumner P, Nachev P, Morris P, Peters AM, Jackson SR, Kennard C, Husain M (2007) Human medial frontal cortex mediates unconscious inhibition of voluntary action. Neuron 54:697–711.

Synofzik M, Thier P, Lindner A (2006) Internalizing agency of self-action: perception of one's own hand movements depends on an adaptable prediction about the sensory action outcome. J Neurophysiol 96:1592–1601.

Synofzik M, Vosgerau G, Newen A (2008) Beyond the comparator model: a multifactorial two-step account of agency. Conscious Cogn 17:219–239.

Takahata K, Takahashi H, Maeda T, Umeda S, Suhara T, Mimura M, Kato M (2012) It's not my fault: postdictive modulation of intentional binding by monetary gains and losses. PloS One 7:e53421.

Thaler D, Chen YC, Nixon PD, Stern CE, Passingham RE (1995) The functions of the medial premotor cortex. I. Simple learned movements. Exp Brain Res 102:445–460.

Thura D, Cisek P (2016) Modulation of Premotor and Primary Motor Cortical Activity during Volitional Adjustments of Speed-Accuracy Trade-Offs. J Neurosci Off J Soc Neurosci 36:938–956.

Tremblay S, Lepage J-F, Latulipe-Loiselle A, Fregni F, Pascual-Leone A, Théoret H (2014) The Uncertain Outcome of Prefrontal tDCS. Brain Stimulat 7:773–783.

Tsakiris M, Longo MR, Haggard P (2010) Having a body versus moving your body: neural signatures of agency and body-ownership. Neuropsychologia 48:2740–2749.

Usher M, McClelland JL (2001) The time course of perceptual choice: the leaky, competing accumulator model. Psychol Rev 108:550–592.

Van Rijn H, Kononowicz TW, Meck WH, Ng KK, Penney TB (2011) Contingent negative variation and its relation to time estimation: a theoretical evaluation. Front Integr Neurosci 5:91.

Verbaarschot C, Farquhar J, Haselager P (2015) Lost in time...: The search for intentions and Readiness Potentials. Conscious Cogn 33C:300–315.

Voss M, Ingram JN, Haggard P, Wolpert DM (2006) Sensorimotor attenuation by central motor command signals in the absence of movement. Nat Neurosci 9:26–27.

Voss M, Moore J, Hauser M, Gallinat J, Heinz A, Haggard P (2010) Altered awareness of action in schizophrenia: a specific deficit in predicting action consequences. Brain J Neurol 133:3104–3112.

Walter WG, Cooper R, Aldridge VJ, McCALLUM WC, Winter AL (1964) Contingent Negative Variation : An Electric Sign of Sensori-Motor Association and Expectancy in the Human Brain. Nature 203:380–384.

Waszak F, Wascher E, Keller P, Koch I, Aschersleben G, Rosenbaum DA, Prinz W (2005) Intention-based and stimulus-based mechanisms in action selection. Exp Brain Res 162:346–356.

Wegner DM (2003a) The Illusion of Conscious Will, New Ed edition. MIT Press.

Wegner DM (2003b) The mind's best trick: how we experience conscious will. Trends Cogn Sci 7:65–69.

Wegner DM, Sparrow B, Winerman L (2004) Vicarious agency: experiencing control over the movements of others. J Pers Soc Psychol 86:838–848.

Wegner DM, Wheatley T (1999) Apparent mental causation. Sources of the experience of will. Am Psychol 54:480–492.

Welchman AE, Stanley J, Schomers MR, Miall RC, Bülthoff HH (2010) The quick and the dead: when reaction beats intention. Proc Biol Sci 277:1667–1674.

Wenke D, Fleming SM, Haggard P (2010) Subliminal priming of actions influences sense of control over effects of action. Cognition 115:26–38.

Wilson B, Cockburn J, Baddeley A (1985) Test di memoria comportamentale Rivermead, tr. it., Firenze: Giunti-Organizzazioni Speciali.

Wittgenstein L (2009) Philosophical Investigations (Hacker PMS, Schulte J, eds)., 4th Edition edition. Chichester, West Sussex, U.K. ; Malden, MA: Wiley-Blackwell.

Wolpe N, Haggard P, Siebner HR, Rowe JB (2013) Cue integration and the perception of action in intentional binding. Exp Brain Res 229:467–474.

Wolpe N, Moore JW, Rae CL, Rittman T, Altena E, Haggard P, Rowe JB (2014) The medial frontal-prefrontal network for altered awareness and control of action in corticobasal syndrome. Brain J Neurol 137:208–220.

Wolpe N, Rowe JB (2014) Beyond the “urge to move”: objective measures for the study of agency in the post-Libet era. Front Hum Neurosci 8:450.

Wolpert DM, Ghahramani Z (2000) Computational principles of movement neuroscience. Nat Neurosci 3 Suppl:1212–1217.

Wolpert DM, Ghahramani Z, Jordan MI (1995) An internal model for sensorimotor integration. Science 269:1880–1882.

Woods AJ, Hamilton RH, Kranjec A, Minhaus P, Bikson M, Yu J, Chatterjee A (2014) Space, time, and causality in the human brain. NeuroImage 92:285–297.

Yoshie M, Haggard P (2013) Negative emotional outcomes attenuate sense of agency over voluntary actions. Curr Biol CB 23:2028–2032.

Zaehle T, Sandmann P, Thorne JD, Jäncke L, Herrmann CS (2011) Transcranial direct current stimulation of the prefrontal cortex modulates working memory performance: combined behavioural and electrophysiological evidence. BMC Neurosci 12:2.

Zama T, Shimada S (2015) Simultaneous measurement of electroencephalography and near-infrared spectroscopy during voluntary motor preparation. Sci Rep 5:16438.

Zimmerman P, Fimm B (1992) Testbatterie zur Aufmerksamkeitsprufung (TAP). In: Wurselen: Psytest.

Zoccolotti P, Pizzamiglio L, Pittau P, Galati G (1994) Batteria di Test per l’esame dell’Attenzione. In Roma: Psytest.

Zwissler B, Sperber C, Aigeldinger S, Schindler S, Kissler J, Plewnia C (2014) Shaping Memory Accuracy by Left Prefrontal Transcranial Direct Current Stimulation. J Neurosci 34:4022–4026.